

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 12:36:27 : Search time 67.61 Seconds
(without alignments)
199.783 Million cell updates/sec

Title: US-09-838-785-2

Perfect score: 2861
Sequence: 1 MVOGLMWSRLHRRKAQLL.....AIYFATGVPEKSDAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Issued Patents: AA.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCYUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------------------------|
| 1 | 2861 | 100.0 | 553 | 4 | US-09-020-956-113 Sequence 113, App |
| 2 | 2861 | 100.0 | 553 | 4 | US-09-030-607-113 Sequence 113, App |
| 3 | 2861 | 100.0 | 553 | 4 | US-09-439-313-113 Sequence 113, App |
| 4 | 1287 | 45.0 | 255 | 4 | US-09-071-710-36 Sequence 36, Appl |
| 5 | 1287 | 45.0 | 255 | 4 | US-09-525-397-36 Sequence 36, Appl |
| 6 | 452 | 15.8 | 84 | 4 | US-09-439-313-571 Sequence 571, App |
| 7 | 330 | 11.5 | 516 | 1 | US-08-356-340-4 Sequence 4, Appl |
| 8 | 330 | 11.5 | 516 | 2 | US-08-786-555-4 Sequence 4, Appl |
| 9 | 324.5 | 11.3 | 525 | 1 | US-08-356-340-2 Sequence 2, Appl |
| 10 | 324.5 | 11.3 | 525 | 2 | US-08-786-555-2 Sequence 2, Appl |
| 11 | 312 | 10.9 | 58 | 4 | US-09-439-313-547 Sequence 547, App |
| 12 | 304 | 10.6 | 56 | 4 | US-09-439-313-564 Sequence 564, App |
| 13 | 243 | 8.5 | 44 | 4 | US-09-071-710-37 Sequence 37, Appl |
| 14 | 243 | 8.5 | 44 | 4 | US-09-525-397-37 Sequence 37, Appl |
| 15 | 157 | 5.5 | 29 | 4 | US-09-439-313-546 Sequence 546, App |
| 16 | 146 | 5.1 | 27 | 4 | US-09-071-710-39 Sequence 39, Appl |
| 17 | 146 | 5.1 | 27 | 4 | US-09-525-397-39 Sequence 39, Appl |
| 18 | 146 | 5.1 | 27 | 4 | US-09-439-313-566 Sequence 566, App |
| 19 | 137 | 4.8 | 25 | 4 | US-09-439-313-520 Sequence 520, Appl |
| 20 | 135 | 4.7 | 27 | 4 | US-09-071-710-38 Sequence 38, Appl |
| 21 | 135 | 4.7 | 27 | 4 | US-09-525-397-38 Sequence 38, Appl |
| 22 | 120 | 4.2 | 21 | 4 | US-09-439-313-496 Sequence 496, App |
| 23 | 120 | 4.2 | 21 | 4 | US-09-439-313-521 Sequence 521, App |
| 24 | 119 | 4.2 | 22 | 4 | US-09-439-313-565 Sequence 565, App |
| 25 | 115 | 4.0 | 22 | 4 | US-09-439-313-559 Sequence 559, App |
| 26 | 114 | 4.0 | 22 | 4 | US-09-439-313-557 Sequence 557, App |
| 27 | 114 | 4.0 | 22 | 4 | US-09-439-313-561 Sequence 561, App |

| | | | | | |
|----|-------|-----|-----|---|-------------------------------------|
| 28 | 112 | 3.9 | 20 | 4 | US-09-439-313-495 Sequence 495, App |
| 29 | 110 | 3.8 | 427 | 4 | US-08-900-230-4 Sequence 4, Appl |
| 30 | 109.5 | 3.8 | 455 | 1 | US-08-035-928-2 Sequence 2, Appl |
| 31 | 109 | 3.8 | 22 | 4 | US-09-439-313-555 Sequence 555, App |
| 32 | 107 | 3.7 | 20 | 4 | US-09-439-313-453 Sequence 493, App |
| 33 | 107 | 3.7 | 20 | 4 | US-09-439-313-497 Sequence 497, App |
| 34 | 106.5 | 3.7 | 436 | 6 | 5432081-10 Patent No. 5432081 |
| 35 | 106 | 3.7 | 22 | 4 | US-09-439-313-569 Sequence 569, App |
| 36 | 106 | 3.7 | 427 | 4 | US-09-198-737-4 Sequence 4, Appl |
| 37 | 105 | 3.7 | 22 | 4 | US-09-439-313-553 Sequence 553, App |
| 38 | 105 | 3.7 | 22 | 4 | US-09-439-313-572 Sequence 572, App |
| 39 | 101 | 3.5 | 18 | 4 | US-09-439-313-548 Sequence 548, App |
| 40 | 100 | 3.5 | 20 | 4 | US-09-439-313-498 Sequence 490, App |
| 41 | 100 | 3.5 | 20 | 4 | US-09-439-313-499 Sequence 499, App |
| 42 | 99 | 3.5 | 20 | 4 | US-09-439-313-553 Sequence 553, App |
| 43 | 99 | 3.5 | 20 | 4 | US-09-439-313-570 Sequence 570, App |
| 44 | 98 | 3.4 | 20 | 4 | US-09-439-313-498 Sequence 498, App |
| 45 | 98 | 3.4 | 713 | 3 | US-09-335-409-11 Sequence 11, Appl |

ALIGNMENTS

RESULT 1
US-09-020-956-113
Sequence 113, Application US/09020956
Patent No. 6261362
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillu, Davin C.
TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
NUMBER OF SEQUENCES: 178
CORRESPONDENCE ADDRESSES:
ADDRESS: SEED AND BERRY LLP
ADDRESS: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 09-FEB-1998
APPLICATION NUMBER: US/09/020,956
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Mark David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.427C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 553 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
US-09-020-956-113

Query Match 100.0% Score 2861: DB 4: Length 553:
Best Local Similarity 100.0% Pred. No. 3.7e-271:
Matches 553: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
1 MVOGLMWSRLHRRKAQLLVLTFGLFVCLAGITVPPPLLEVGVEKEMTVNLGIG 60
|||||

```

Db      1  MVQRLMWSLTLHRRKAQOLLNVNLTFFGLEVCIAAGITTYVPLLLEVGVEKEFTWVLG16 60
OY      61  PYLGLVCPVLLGSADSHMWRGRRRPFIMWLSLGLLSLFLIPRGMWLAGLCCPPRPL 120
Db      61  PYLGLVCPVLLGSADSHMWRGRRRPFIMWLSLGLLSLFLIPRGMWLAGLCCPPRPL 120
OY      121  ELTALLILGGLDFGGOVCFPLEALLSDLFRDPDHCROAVSYAAMISLGGCLGYLLPA 180
Db      121  ELTALLILGGLDFGGOVCFPLEALLSDLFRDPDHCROAVSYAAMISLGGCLGYLLPA 180
OY      181  IDMDTSALAPYLGTQOECLFGLLTLLFLTCVAAATLLVAEAAIGPTEPAEGISAPLSPH 240
Db      181  IDMDTSALAPYLGTQOECLFGLLTLLFLTCVAAATLLVAEAAIGPTEPAEGISAPLSPH 240
OY      241  CCPCRARLAFNRUGALLPRLHOLCCMRPTLRLPFAELCSMMALMTFFLTDFPEBGL 300
Db      241  CCPCRARLAFNRUGALLPRLHOLCCMRPTLRLPFAELCSMMALMTFFLTDFPEBGL 300
OY      301  YQGVRAEPGRTARRHYDGVNRGSLGLFLQCAISLVFSVMDRLVQRGTBRAVYLA5VA 360
Db      301  YQGVRAEPGRTARRHYDGVNRGSLGLFLQCAISLVFSVMDRLVQRGTBRAVYLA5VA 360
OY      361  APVVAAGATCLISHVAVVTASAALTGTFPSALQILEPTLASLYHREKQVFLPKYRGDTG 420
Db      361  APVVAAGATCLISHVAVVTASAALTGTFPSALQILEPTLASLYHREKQVFLPKYRGDTG 420
OY      421  ASSEDSLMTSLFPGKPGAPFPGNHVAGAGSGCLPPPALCGASACDVSVRVVVGEPTEA 480
Db      421  ASSEDSLMTSLFPGKPGAPFPGNHVAGAGSGSGCLPPPALCGASACDVSVRVVVGEPTEA 480
OY      481  RVVPGGICITLDLAILDSAPLLSOVAPSLFMGSIYVOLSQSVTAMVSAAGIYAFATQ 540
Db      481  RVVPGGICITLDLAILDSAPLLSOVAPSLFMGSIYVOLSQSVTAMVSAAGIYAFATQ 540
OY      541  VVFEKSDILAKYSA 553
Db      541  VVFEKSDILAKYSA 553

RESULT 2
US-09-030-607-113
: Sequence 113, Application US/09030607
: Patent No. 6262245
: GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun
: APPLICANT: Dillon, Davin C.
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FOR
: NUMBER OF SEQUENCES: 224
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: SEED and BERRY LLP
: STREET: 6300 Columbia Center, 701 Fifth Avenue
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98104
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/030,607
: FILING DATE: 25-FEB-1998
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: Maki, David J.
: REGISTRATION NUMBER: 31,392
: REFERENCE/DOCKET NUMBER: 210121.427C3
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 622-4900
: TELEFAX: (206) 682-6031
: INFORMATION FOR SEQ ID NO: 113:
: SEQUENCE CHARACTERISTICS:

```

```

; LENGTH 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-09-030-607-113

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 3,76-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MQRILWVSLRHRKRAQLLVNLITFGLGVCLAAGITTPYRLLEKGVEEKEMTWLIG 60
DB 1 MQRILWVSLRHRKRAQLLVNLITFGLGVCLAAGITTPYRLLEKGVEEKEMTWLIG 60
QY 61 PVLGLVCPVLDSASPDHMGGRYGRRRPFIMWLSIGLLSLFLIPRGMWLAGLCCDP 120
DB 61 PVLGLVCPVLDSASPDHMGGRYGRRRPFIMWLSIGLLSLFLIPRGMWLAGLCCDP 120
QY 121 ELALLILGVGLDFCGQVCFTEPLLEALLSDLFDPDHCRAQSVYAAEMISAGCLGYL 180
DB 121 ELALLILGVGLDFCGQVCFTEPLLEALLSDLFDPDHCRAQSVYAAEMISAGCLGYL 180
QY 181 IDWDTSALAPYLGTQBECLFGLITFLITCVAAATLVAEALGTEPAEGLSASLS 240
DB 181 IDWDTSALAPYLGTQBECLFGLITFLITCVAAATLVAEALGTEPAEGLSASLS 240
QY 241 GCPGRARLAFRNLGALPRLHOLCCMPRTLRLPYAELCSMMALMTFTLFTDFV 300
DB 241 GCPGRARLAFRNLGALPRLHOLCCMPRTLRLPYAELCSMMALMTFTLFTDFV 300
QY 301 YGVGPRAEPTGARRRHYDEGVRMGSLGLFLQCAISLVSFLVMDRLVORPGR 360
DB 301 YGVGPRAEPTGARRRHYDEGVRMGSLGLFLQCAISLVSFLVMDRLVORPGR 360
QY 361 APFVAAGATCLSHVAVVTASALIGTFTESAQLILPYTASLYHREKQVFLPK 420
DB 361 APFVAAGATCLSHVAVVTASALIGTFTESAQLILPYTASLYHREKQVFLPK 420
QY 421 ASSEDSLMTSFLPGRPGAPFPRNGHAGAGSGILPPRLCGASACDVSVRVV 480
DB 421 ASSEDSLMTSFLPGRPGAPFPRNGHAGAGSGILPPRLCGASACDVSVRVV 480
QY 481 RVPVPGICLDAIILDSAPLISOVAPSLFMGSIYQLDSQSVTAYMVAAGL 540
DB 481 RVPVPGICLDAIILDSAPLISOVAPSLFMGSIYQLDSQSVTAYMVAAGL 540
QY 541 VVEDKSDLAKYSA 553
DB 541 VVEDKSDLAKYSA 553

RESULT 3
US-09-439-313-113
; Sequence 113, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yugu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solt, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER

```

FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439.313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 113
LENGTH: 553
TYPE: PRT
ORGANISM: Homo sapien
US-09-439-313-113

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVQRLVWSRLHRRKAQOLLVNLTFEGLEVCLAAGITVPPPLLEVEGEKEMTWLGIG 60
DB 1 MVQRLVWSRLHRRKAQOLLVNLTFEGLEVCLAAGITVPPPLLEVEGEKEMTWLGIG 60
QY 61 PVLGLVCPVLGASDHRGRRRPTWALSGLISLFLIPRAGWLAGLCPDPRPL 120
DB 61 PVLGLVCPVLGASDHRGRRRPTWALSGLISLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLIILGVLDFGGOVCFPLLEALLSDFRDPHCRQAVSVAFMISLGGCIGLILPA 180
DB 121 ELALLIILGVLDFGGOVCFPLLEALLSDFRDPHCRQAVSVAFMISLGGCIGLILPA 180
QY 181 IDMDTSALAPYLGTQECIFGLTLIFLTCVAATLLVAEALGFTPEAGLSAPSLSPH 240
DB 181 IDMDTSALAPYLGTQECIFGLTLIFLTCVAATLLVAEALGFTPEAGLSAPSLSPH 240
QY 241 IDMDTSALAPYLGTQECIFGLTLIFLTCVAATLLVAEALGFTPEAGLSAPSLSPH 240
DB 241 IDMDTSALAPYLGTQECIFGLTLIFLTCVAATLLVAEALGFTPEAGLSAPSLSPH 240
QY 241 CCEPCARLAFRLGALPRLHOLCCMRPTLRRLFAELCSMALMTFLFTDVEGEL 300
DB 241 CCEPCARLAFRLGALPRLHOLCCMRPTLRRLFAELCSMALMTFLFTDVEGEL 300
QY 301 YGVGPRAEPTGARHRYDEGVMSGLFLQCAISLVFSLVMDRLVORGTAVYLA 360
DB 301 YGVGPRAEPTGARHRYDEGVMSGLFLQCAISLVFSLVMDRLVORGTAVYLA 360
QY 361 AEPVAAAGATCISHVAVVTASALGTFTSALQILPYTLASLYHREKOVFLPKYRGDTGG 420
DB 361 AEPVAAAGATCISHVAVVTASALGTFTSALQILPYTLASLYHREKOVFLPKYRGDTGG 420
QY 421 ASSESLSMTSFLPGPKPAPFPNGHVAGAGSGLPPPALCGASACDVSVRYVGEPT 480
DB 421 ASSESLSMTSFLPGPKPAPFPNGHVAGAGSGLPPPALCGASACDVSVRYVGEPT 480
QY 481 RVPVGRGICLDLALILDSAFLLSQVAPSLFMGSIYVLSQSVTAYVMSAAGLIVAIYFA 540
DB 481 RVPVGRGICLDLALILDSAFLLSQVAPSLFMGSIYVLSQSVTAYVMSAAGLIVAIYFA 540
QY 541 VFEDKSDLA 553
DB 541 VFEDKSDLA 553

RESULT 4
US-09-071-710-36
Sequence 36, Application US/09071710
Patent No. 6130043
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANDOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071.710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850.713
FILING DATE: 02-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083. US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6130043e
US-09-071-710-36

Query Match 45.0%; Score 1287; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 8.1e-118;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GLYGVPRAEPTGARHRYDEGVMSGLFLQCAISLVFSLVMDRLVORGTAVYLA 358
DB 1 GLYGVPRAEPTGARHRYDEGVMSGLFLQCAISLVFSLVMDRLVORGTAVYLA 358
QY 359 VAFVPAAGATCISHVAVVTASALGTFTSALQILPYTLASLYHREKOVFLPKYRGDT 418
DB 61 VAFVPAAGATCISHVAVVTASALGTFTSALQILPYTLASLYHREKOVFLPKYRGDT 418
QY 419 GGASSEDLSMTSFLPGPKPAPFPNGHVAGAGSGLPPPALCGASACDVSVRYVGEPT 478
DB 121 GGASSEDLSMTSFLPGPKPAPFPNGHVAGAGSGLPPPALCGASACDVSVRYVGEPT 478
QY 479 EARVVPGRGICLDLALILDSAFLLSQVAPSLFMGSIYVLSQSVTAYVMSAAGLIVAIYFA 538
DB 181 EARVVPGRGICLDLALILDSAFLLSQVAPSLFMGSIYVLSQSVTAYVMSAAGLIVAIYFA 538
QY 539 TVGVFEDKSDLA 553
DB 241 TVGVFEDKSDLA 553

RESULT 5
US-09-525-397-36
Sequence 36, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULIA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KASS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: IBM Compatible
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525,397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083-US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-36

Query Match 45.0%; Score 1287; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 8.1e-118;
Matches 255; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 1 GYGVPRAEGETARRHYDEGVKMSGLFLCAISLVESLMDRLVQREGTRAVYLAS 60
QY 299 GYGVPRAEGETARRHYDEGVKMSGLFLCAISLVESLMDRLVQREGTRAVYLAS 358
Db 1 GYGVPRAEGETARRHYDEGVKMSGLFLCAISLVESLMDRLVQREGTRAVYLAS 60
QY 359 VAAPVAGATCLSHSVAVVTASALNGFTPSALQIIPYTLASLYHREKOVFLPYRGDT 418
Db 61 VAAPVAGATCLSHSVAVVTASALNGFTPSALQIIPYTLASLYHREKOVFLPYRGDT 120
QY 419 GGASEDSLMTSFLPGPKGAPFPNGHVGAGSGILLPPPALCGASACDVSVRVVGEPT 478
Db 121 GGASEDSLMTSFLPGPKGAPFPNGHVGAGSGILLPPPALCGASACDVSVRVVGEPT 180
QY 479 EARVPGGICLDAIIDSATLLSOVAPSLMGSLVQISQSVTAVMSAAGLGLVAITFA 538
Db 181 EARVPGGICLDAIIDSATLLSOVAPSLMGSLVQISQSVTAVMSAAGLGLVAITFA 240
QY 539 TOVFDKSDLAKYSA 553
Db 241 TOVFDKSDLAKYSA 255

RESULT 6
US-09-439-313-571
Sequence 571, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yugu
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427c9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO: 571
LENGTH: 84
TYPE: PRT
ORGANISM: Homo sapiens
US-09-439-313-571

Query Match 15.8%; Score 452; DB 4; Length 84;
Best Local Similarity 100.0%; Pred. No. 6.5e-37;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 404 HREKOVFLPYRGDTGAGSSDSLMTSFLPGPKGAPFPNGHVGAGSGILLPPPALCGA 463
Db 1 HREKOVFLPYRGDTGAGSSDSLMTSFLPGPKGAPFPNGHVGAGSGILLPPPALCGA 60
QY 464 SAGDSVRYVVGEPTEARVYVGRG 487
Db 61 SAGDSVRYVVGEPTEARVYVGRG 84

RESULT 7
US-08-356-340-4
Sequence 4, Application US/08356340
Patent No. 5608146
GENERAL INFORMATION:
APPLICANT: FROMMER, Wolf-Bernd
APPLICANT: RIESMEIER, Jorg
TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE
TITLE OF INVENTION: TRANSPORTER, PLASMIDS, BACTERIA AND PLANTS CONTAINING A
TITLE OF INVENTION: TRANSPORTER AS WELL AS A PROCESS FOR THE PREPARATION AND
TITLE OF INVENTION: TRANSFORMATION OF YEAST STRAINS FOR THE IDENTIFI...
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ostrolenk, Faber, Gerb & Soffen
STREET: 1180 Avenue of the Americas
CITY: New York
STATE: NY
COUNTRY: US
ZIP: 10036-8403
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/356,340
FILING DATE: 21-DEC-1994
CLASSIFICATION: 800


```
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US PCT/EP93/01604
; FILING DATE: 22-JUN-1993
; PRIOR APPLICATION DATA: DE P42 20 759.2
; APPLICATION NUMBER: DE P42 20 759.2
; FILING DATE: 24-JUN-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Meilman, Edward A.
; REGISTRATION NUMBER: 24,735
; REFERENCE/DOCKET NUMBER: P/951-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 382-0700
; TELEFAX: (212) 382-0888
; TELEX: 236925
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-356-340-4

Query Match      11.5%; Score 330; DB 1; Length 516;
Best Local Similarity 24.7%; Pred. No. 7.4e-24;
Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

OY 4 RLWVRRLRHRKAQOLLVNLTFGLVCLAGITVYPPILLEVEVEEKKFMVIGIGPVL 63
   29 KIM-----KIIVASIAAGVQFGWALDLSLTPYVQLGIPHKFASFIWLCGPIS 78
OY 64 GLVCPVLGASDHWGRGRYRRPITWALSIGILSLFLIPRAGWLAGL-----LCPD 116
   79 GMIVQPVVGVYSDNCSSRFGRRRPFIAGALVMTAVFLI---GRAADLGHASGDTLKG 135
OY 117 PRRLALLILGVLDDCGVCFPLEALLSDLEFRDPD-HCROAVSYAAMISLGGCLG 175
   136 FKPRAIIVVGVFWILDVANNMLGQPCRLADLSGKSGRRRTANAFSFFMAVGNILG 195
OY 176 YLLPAIDW-----DTSALAPYIGTOECLF-GLTLIFLTCAATILVAEE-AALGP 225
   196 YAAGSYSHLFKVPFSKTRKACMYCANLKSCEFFIAIFLLSLTTALTILVRENELPEKDE 255
OY 226 TERPAGLSAPSLSPHCCPCRRARLAFRNIGALLPRLHLOCCRMPTLRRLFEVLECSMAL 285
   256 QEIDKRLGAG-----KSKVPE--FGEIRGALKEL---PRPMWILLVTCLNMTIAM 301
OY 286 MTFLEFYDPEVGEGLYGVYPRAPGTEARRHYDEGVRMGSLGLFQCAISLVSIVMDRL 345
   302 PFFFLYDIDDMAKEVFGG---QVGD--ARLYDLGVRAAGMLLLQSVVLGFMSLGVEFL 355
OY 346 VORFGTRAVYLASVAAPVAAGATCLSHSVAVY-----TASALLTGFT----- 388
   366 GKRIIG-GAKRILMGIINFVL---AICLAMTILVTKMAEKSRQHPDAGTLMGPTGVKIGAL 411
OY 389 --FSALQI-----LPYTLASLYHREK-----QVFLPKYRGDTGASSEDSS 426
   412 LLFALGIPLATFISIPALASIFSSNRSGGSLGVNLINAIIVPQMLVSLVGPMDL 471
OY 427 LMTSFLPGKPRKAPFPNGHVAGSGG-----LLPPPALCGASACDVSRYVYVG 475
   472 FCGGNLPG-----FVVGAVAAAASAVLALITMLPSPPA-----DARPAVAMG 512
DB 472 FCGGNLPG-----FVVGAVAAAASAVLALITMLPSPPA-----DARPAVAMG 512

RESULT      8
; US-08-786-555-4
; Sequence 4, Application US/08786555B
; Patent No. 5981181
; GENERAL INFORMATION:
; APPLICANT: FROMMER, WOLF-BERND
; APPLICANT: RIESMEIER, JORG
; TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE TRANSPORTER,
; TITLE OF INVENTION: PLASMIDS, BACTERIA AND PLANTS CONTAINING A TRANSPORTER
```

```
; TITLE OF INVENTION: AS WELL AS A PROCESS FOR PREPARATION AND TRANSFORMATION
; TITLE OF INVENTION: OF YEAST STRAINS FOR THE IDENTIFICATION
; FILE REFERENCE: FROMMER
; CURRENT APPLICATION NUMBER: US/08/786,555B
; CURRENT FILING DATE: 1997-01-21
; EARLIER APPLICATION NUMBER: 08/356,340
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: PCT/EP93/01604
; EARLIER FILING DATE: 1993-06-22
; EARLIER APPLICATION NUMBER: DE P4220759.2
; EARLIER FILING DATE: 1992-06-24
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 4
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Spinacia oleracea
; US-08-786-555-4

Query Match      11.5%; Score 330; DB 2; Length 516;
Best Local Similarity 24.7%; Pred. No. 7.4e-24;
Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

OY 4 RLWVRRLRHRKAQOLLVNLTFGLVCLAGITVYPPILLEVEVEEKKFMVIGIGPVL 63
   29 KIM-----KIIVASIAAGVQFGWALDLSLTPYVQLGIPHKFASFIWLCGPIS 78
OY 64 GLVCPVLGASDHWGRGRYRRPITWALSIGILSLFLIPRAGWLAGL-----LCPD 116
   79 GMIVQPVVGVYSDNCSSRFGRRRPFIAGALVMTAVFLI---GRAADLGHASGDTLKG 135
OY 117 PRRLALLILGVLDDCGVCFPLEALLSDLEFRDPD-HCROAVSYAAMISLGGCLG 175
   136 FKPRAIIVVGVFWILDVANNMLGQPCRLADLSGKSGRRRTANAFSFFMAVGNILG 195
OY 176 YLLPAIDW-----DTSALAPYIGTOECLF-GLTLIFLTCAATILVAEE-AALGP 225
   196 YAAGSYSHLFKVPFSKTRKACMYCANLKSCEFFIAIFLLSLTTALTILVRENELPEKDE 255
OY 226 TERPAGLSAPSLSPHCCPCRRARLAFRNIGALLPRLHLOCCRMPTLRRLFEVLECSMAL 285
   256 QEIDKRLGAG-----KSKVPE--FGEIRGALKEL---PRPMWILLVTCLNMTIAM 301
OY 286 MTFLEFYDPEVGEGLYGVYPRAPGTEARRHYDEGVRMGSLGLFQCAISLVSIVMDRL 345
   302 PFFFLYDIDDMAKEVFGG---QVGD--ARLYDLGVRAAGMLLLQSVVLGFMSLGVEFL 355
OY 346 VORFGTRAVYLASVAAPVAAGATCLSHSVAVY-----TASALLTGFT----- 388
   366 GKRIIG-GAKRILMGIINFVL---AICLAMTILVTKMAEKSRQHPDAGTLMGPTGVKIGAL 411
OY 389 --FSALQI-----LPYTLASLYHREK-----QVFLPKYRGDTGASSEDSS 426
   412 LLFALGIPLATFISIPALASIFSSNRSGGSLGVNLINAIIVPQMLVSLVGPMDL 471
OY 427 LMTSFLPGKPRKAPFPNGHVAGSGG-----LLPPPALCGASACDVSRYVYVG 475
   472 FCGGNLPG-----FVVGAVAAAASAVLALITMLPSPPA-----DARPAVAMG 512
DB 472 FCGGNLPG-----FVVGAVAAAASAVLALITMLPSPPA-----DARPAVAMG 512

RESULT      9
; US-08-356-340-2
; Sequence 2, Application US/08356340
; Patent No. 5608146
; GENERAL INFORMATION:
; APPLICANT: FROMMER, WOLF-BERND
; APPLICANT: RIESMEIER, JORG
; TITLE OF INVENTION: DNA SEQUENCES WITH OLIGOSACCHARIDE
; TITLE OF INVENTION: TRANSPORTER, PLASMIDS, BACTERIA AND PLANTS CONTAINING A
; TITLE OF INVENTION: TRANSPORTER AS WELL AS A PROCESS FOR THE PREPARATION AND
; TITLE OF INVENTION: TRANSFORMATION OF YEAST STRAINS FOR THE IDENTIFI...
```


[illegible]

```

: SOFTWARE: FASTSEQ for Windows Version 3.0
: SEQ ID NO 564
: LENGTH: 56
: TYPE: PRT
: ORGANISM: Homo sapiens
US-09-439-313-564

Query Match          10.6%: Score 304; DB 4; Length 56;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 56: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      218 AEEALGTFEEAGHSAPSLSPHCCPCARLAFRNLGALLPRLHQLCCRMPTLRR 273
          |||||||.....
Db       1 AEEALGTFEEAGHSAPSLSPHCCPCARLAFRNLGALLPRLHQLCCRMPTLRR 56

RESULT 13
US-09-071-710-37
: Sequence 37, Application us/09071710
: Patent No. 6130043
: GENERAL INFORMATION:
: APPLICANT: BILLING-MEDEL, PATRICIA
: APPLICANT: COHEN, MAURICE
: APPLICANT: COLPITTS, TRACEY L.
: APPLICANT: FRIEDMAN, PAULA N.
: APPLICANT: GORDON, JULIAN
: APPLICANT: GRANADOS, EDWARD N.
: APPLICANT: HODGES, STEVEN C.
: APPLICANT: KLAAS, MICHAEL R.
: APPLICANT: KRATOCHVIL, JON D.
: APPLICANT: ROBERTS-RAPP, LISA
: APPLICANT: RUSSELL, JOHN C.
: APPLICANT: STROUPE, STEPHEN D.
: TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
: TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
: NUMBER OF SEQUENCES: 41
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Abbott Laboratories
: STREET: 100 Abbott Park Road
: CITY: Abbott Park
: STATE: IL
: COUNTRY: USA
: ZIP: 60064-3500
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/071,710
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/850,713
: FILING DATE: 02-MAY-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Becker, Cheryl L.
: REGISTRATION NUMBER: 35,441
: REFERENCE/DOCKET NUMBER: 6083_US.P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 847/935-1729
: TELEFAX: 847/938-2623
: TELEX:
: INFORMATION FOR SEQ ID NO: 37:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 44 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: No. 6130043e
US-09-071-710-37

```

Query Match 8.5%; Score 243; DB 4; Length 44;
Best Local Similarity 100.0%; Pred. No. 6.8e-17;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 403 YHREKQVFLPKRYGDTGASSEDLSMTSFLPGPKGAPFPNGHV 446
Db 1 YHREKQVFLPKRYGDTGASSEDLSMTSFLPGPKGAPFPNGHV 44

RESULT 14

US-09-525-397-37
; Sequence 37, Application US/09525397
; Patent No. 6252047

GENERAL INFORMATION:

APPLICANT: BILLING-MEDEL, PATRICIA

APPLICANT: COHEN, MAURICE

APPLICANT: COLPITTS, TRACEY L.

APPLICANT: FRIEDMAN, PAULA N.

APPLICANT: GORDON, JULIAN

APPLICANT: GRANADOS, EDWARD N.

APPLICANT: HODGES, STEVEN C.

APPLICANT: KLAS, MICHAEL R.

APPLICANT: KRATOCHVIL, JON D.

APPLICANT: ROBERTS-RAPP, LISA

APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.

TITLE OF INVENTION: REAGENTS AND METHODS USEFUL

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Abbott Laboratories

STREET: 100 Abbott Park Road

CITY: Abbott Park

STATE: IL

COUNTRY: USA

ZIP: 60064-3500

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/525,397

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/071,710

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Becker, Cheryl L.

REGISTRATION NUMBER: 35,441

REFERENCE/DOCKET NUMBER: 6083.US.PI

TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729

TELEFAX: 847/938-2623

TELEX:

INFORMATION FOR SEQ ID NO: 37:

SEQUENCE CHARACTERISTICS:

LENGTH: 44 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: No. 6252047e

US-09-525-397-37

Query Match 8.5%; Score 243; DB 4; Length 44;

Best Local Similarity 100.0%; Pred. No. 6.8e-17;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 403 YHREKQVFLPKRYGDTGASSEDLSMTSFLPGPKGAPFPNGHV 446

Db 1 YHREKQVFLPKRYGDTGASSEDLSMTSFLPGPKGAPFPNGHV 44

RESULT 15

US-09-439-313-546

; Sequence 546, Application US/09439313

; Patent No. 6329505

GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.

APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan Louise

APPLICANT: Jiang Yuqi

APPLICANT: Reed, Steven G.

APPLICANT: Kalos, Michael

APPLICANT: Fanger, Gary

APPLICANT: Retter, Mark

APPLICANT: Solk, John

APPLICANT: Day, Craig

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND

FILE REFERENCE: 210121.427C9

CURRENT APPLICATION NUMBER: US/09/439,313

CURRENT FILING DATE: 1999-11-12

NUMBER OF SEQ ID NOS: 575

SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 546

LENGTH: 29

TYPE: PRT

ORGANISM: Homo sapiens

US-09-439-313-546

Query Match 5.5%; Score 157; DB 4; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.5e-09;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 295 FVGEGLYQGVPRAPGTEARRHYDEGYRM 323
Db 1 FVGEGLYQGVPRAPGTEARRHYDEGYRM 29

Search completed: June 26, 2002, 13:44:21
Job time: 4074 sec

Thu Jun 27 09:46:43 2002

us-09-838-785-2.rai

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 12:34:42 ; Search time 80.68 Seconds
(without alignments)
761.326 Million cell updates/sec

Title: US-09-838-785-2

Perfect score: 2861
Sequence: 1 MVQRLWVSRLRRKQAQLL.....AIYFATQVFDKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_032802:*

- 1: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 2861 | 100.0 | 553 | 19 | AAW71869 |
| 2 | 2861 | 100.0 | 553 | 19 | AAW69385 |
| 3 | 2861 | 100.0 | 553 | 21 | AAW69385 |
| 4 | 2861 | 100.0 | 553 | 21 | AAW69385 |
| 5 | 2861 | 100.0 | 553 | 22 | AAU69763 |
| 6 | 2861 | 100.0 | 553 | 22 | AAU04961 |
| 7 | 2861 | 100.0 | 553 | 22 | AAU01117 |
| 8 | 2861 | 100.0 | 553 | 22 | AAU01117 |
| 9 | 2861 | 100.0 | 553 | 22 | AAU01117 |
| 10 | 2861 | 100.0 | 553 | 22 | AAU01117 |
| 11 | 2861 | 100.0 | 553 | 22 | AAU01117 |

| | | | | | |
|----|--------|------|-----|----|----------|
| 12 | 2596 | 90.7 | 595 | 22 | AAU01318 |
| 13 | 1417.5 | 49.5 | 371 | 22 | AAU69875 |
| 14 | 1417.5 | 49.5 | 371 | 22 | AAU01230 |
| 15 | 1416 | 49.5 | 371 | 22 | AAE01362 |
| 16 | 1403.5 | 49.1 | 400 | 22 | AAU69907 |
| 17 | 1403.5 | 49.1 | 400 | 22 | AAU01262 |
| 18 | 1287 | 45.0 | 255 | 20 | AAU69068 |
| 19 | 1287 | 45.0 | 255 | 21 | AAU69068 |
| 20 | 1287 | 45.0 | 255 | 22 | AAU04205 |
| 21 | 1150 | 40.2 | 231 | 21 | AAU54369 |
| 22 | 1120 | 39.1 | 252 | 22 | AAE01423 |
| 23 | 519.5 | 18.2 | 599 | 22 | ABU60709 |
| 24 | 457.5 | 16.0 | 748 | 22 | AAU01227 |
| 25 | 426 | 14.9 | 123 | 22 | AAU69873 |
| 26 | 426 | 14.9 | 123 | 22 | AAU01228 |
| 27 | 357 | 12.5 | 194 | 20 | AAU41119 |
| 28 | 337.5 | 11.8 | 523 | 20 | AAU41129 |
| 29 | 330 | 11.5 | 515 | 22 | AAU41129 |
| 30 | 328.5 | 11.5 | 594 | 22 | AAU60012 |
| 31 | 326.5 | 11.4 | 667 | 20 | AAU41118 |
| 32 | 325 | 11.4 | 563 | 20 | AAU41125 |
| 33 | 323.5 | 11.3 | 507 | 21 | AAU30206 |
| 34 | 323.5 | 11.3 | 512 | 21 | AAU30205 |
| 35 | 323.5 | 11.3 | 539 | 21 | AAU30204 |
| 36 | 313.5 | 11.0 | 605 | 22 | AAU60018 |
| 37 | 313 | 10.9 | 497 | 20 | AAU41116 |
| 38 | 312 | 10.9 | 58 | 22 | AAU69833 |
| 39 | 312 | 10.9 | 58 | 22 | AAU01188 |
| 40 | 311 | 10.9 | 58 | 22 | AAU69073 |
| 41 | 311 | 10.9 | 533 | 20 | AAU41128 |
| 42 | 307.5 | 10.7 | 500 | 22 | AAU60011 |
| 43 | 307.5 | 10.7 | 500 | 22 | AAU60017 |
| 44 | 305.5 | 10.7 | 501 | 20 | AAU41126 |
| 45 | 303.5 | 10.6 | 519 | 20 | AAU41114 |

ALIGNMENTS

| | |
|----------|---|
| RESULT | 1 |
| AAW71869 | AAW71869 standard; Protein; 553 AA. |
| XX | AAW71869; |
| XX | 06-JAN-1999 (first entry) |
| DT | Amino acid encoded by prostate tumour clone L1-12. |
| XX | Prostate; cancer; tumour; vaccine; immunogen; clone. |
| XX | Homo sapiens. |
| OS | W09837093-A2. |
| XX | 27-AUG-1998. |
| XX | 25-FEB-1998; 98WO-US03492. |
| XX | 09-FEB-1998; 98US-0020956. |
| XX | 25-FEB-1997; 97US-0806099. |
| XX | 01-AUG-1997; 97US-0904804. |
| XX | (CORI-) CORIXA CORP. |
| XX | Dillon DC, Xu J; |
| XX | WPI: 1998-609886/51. |
| XX | N-PSDB; AAV61201. |
| XX | Polypeptides comprising immunogenic portions of prostate proteins - |
| XX | used in a vaccine for the treatment of prostate cancer |

Alpha prepro-P501S
Human prostate CDN
P535 splice varia
Human gene 11 enco
Human prostate pro
Ra12-P501S-E2 cons
Protein encoded by
Human prostate-rel
Prostate-specific
Amino acid sequenc
Human secreted pro
Drosophila melanog
Human polypeptide
Human prostate CDN
P535 splice varia
Soybean sucrose tr
Vicia faba sucrose
S. tuberosum SUT1
A. thaliana SUT2 p
Rice sucrose trans
Wheat sucrose tran
Arabidopsis thalia
Arabidopsis thalia
Arabidopsis thalia
S. tuberosum SUT2
Corn sucrose trans
Human prostate pro
Human prostate-spe
Human prostate-spe
Ricinus communis s
L. esculentum SUT4
S. tuberosum SUT4
Daucus carota sucr
Corn sucrose trans

PS Example 1: Page 82-84; 130pp; English.

XX The present sequence is an immunogenic portion of a prostate tumour
CC protein. The immunogen, or the DNA encoding it, can be used as a
CC vaccine for the treatment of prostate cancer. The immunogen was
CC isolated from a prostate tumour cDNA library obtained by subtracting
CC a prostate tumour cDNA expression library with a normal tissue cDNA
CC library.

XX Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 19; Length 553;

Best Local Similarity 100.0%; Pred. No. 1.8e-269;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRHRKAOLLVNLLTFGLEVCLANGITYVPPLLLEVGEKFMVIGIG 60
DB 1 mvgrlwsrllrhkqaqllvnlitfglevclaagicyppllilevgeekfmvligig 60
QY 61 PVGLVCPVLIGSADHMRGKRGRRRPTWALSIGILSLFLIRAGLALCPDPRPL 120
DB 61 pvlglvcpvlignsadsdhmrgrgyrrprfwalsigilslfliragwlaqlcpdprpl 120
QY 121 ELALLIGVGLIDFCGCGVCFPTLEALSDLEFRDPDHCQAYSVAFMISLGCGLYLPA 180
DB 121 elalllvgvllidfcgvcfptleallsdlefrdpdchcraqsvyafmislgcgllylpa 180
QY 181 IDWTSALAPYLGTQECFLGLTLIFLTCVAATLVAAEALGTPEDAEGLSAPSLSPH 240
DB 181 idwtsalapylgtececlfglltliftcvaatlvaeeaalgtpeaeglspslsph 240
QY 241 CPCRARLAFRNGLALPRLHQLCCRMPTLRRLFVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 cpcrarlafnlgallprlhqccrmptlrrlrfvaeelcswmlmtlftlftdyvgegl 300
QY 301 YQGVRAEPGTEARRHYDEGRMGSLGLFLQCAISLVESLVMDRLVOREGTRAVYLASVA 360
DB 301 yqgvraepgtearrhydegvmgsllglflqcaislveslvmdrlyvregtravylasva 360
QY 361 AFPVAGATCISHSVAAVVTASAAALTGFTFSALQIIPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 afpvaagatcishsvaavvtasaaaltgftfsalqilpytlaalyhrekqvflpkrygdtgg 420
QY 421 ASSEDSLMTSFLPGPKPGAPPNNGHVAGSGSLPPPALCGASCDVSVRVVGEPTPEA 480
DB 421 assedslmtsfllpgpkpgapfnnghvagsgslppppalcgascdsvrvvgeptpea 480
QY 481 RVVPGRGICLDLALIDSAPFLSQVAPSLFMGSIYVLSOSVTAVMVSAAGLIVATYFAAQ 540
DB 481 rvvpgrgicldalidsafllsqvapslfmgsivqlsgsvtcaymvsaaglgivatyfatq 540
QY 541 VVFKSDIAKYSA 553
DB 541 vvfksdiakysa 553

RESULT 2

AAW69385 AAW69385 standard; Protein; 553 AA.

XX AAW69385;

XX 08-DEC-1998 (first entry)

DE prostate tumour specific gene clone LI-12 protein.

XX Prostate tumour specific gene; human; prostate cancer; detection;

XX therapy.

XX Homo sapiens.

XX OS W09837418-A2.

XX 27-AUG-1998.

PD 25-FEB-1998; 98MO-US03690.

XX 09-FEB-1998; 98US-0904809.

PR 25-FEB-1997; 97US-0806596.

PR 01-AUG-1997; 97US-0904809.

XX (CORI-) CORIXA CORP.

XX DILLON DC, Xu J;

XX WPI; 1998-480805/41.

DR N-PDB; AAV58586.

XX Novel human prostate specific tumour protein and fragments - useful
PT for detecting and treating prostate cancers

PS Example 1: Page 87-89; 141pp; English.
XX This sequence is encoded by a human prostate tumour specific gene, and
CC can be used in the method of the invention. The method is for detecting
CC prostate cancer comprises contacting a biological sample with an agent
CC able to bind an immunogenic portion of a prostate protein (such as
CC this protein sequence). An antibody which binds to an immunogenic
CC portion of the prostate protein, and the method can be used to detect,
CC monitor progression of, or treat prostate cancers. The antibody may
CC also be conjugated to a therapeutic agent for use in therapy of prostate
CC cancers.

XX Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 19; Length 553;

Best Local Similarity 100.0%; Pred. No. 1.8e-269;

Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRHRKAOLLVNLLTFGLEVCLANGITYVPPLLLEVGEKFMVIGIG 60
DB 1 mvgrlwsrllrhkqaqllvnlitfglevclaagicyppllilevgeekfmvligig 60
QY 61 PVGLVCPVLIGSADHMRGKRGRRRPTWALSIGILSLFLIRAGLALCPDPRPL 120
DB 61 pvlglvcpvlignsadsdhmrgrgyrrprfwalsigilslfliragwlaqlcpdprpl 120
QY 121 ELALLIGVGLIDFCGCGVCFPTLEALSDLEFRDPDHCQAYSVAFMISLGCGLYLPA 180
DB 121 elalllvgvllidfcgvcfptleallsdlefrdpdchcraqsvyafmislgcgllylpa 180
QY 181 IDWTSALAPYLGTQECFLGLTLIFLTCVAATLVAAEALGTPEDAEGLSAPSLSPH 240
DB 181 idwtsalapylgtececlfglltliftcvaatlvaeeaalgtpeaeglspslsph 240
QY 241 CPCRARLAFRNGLALPRLHQLCCRMPTLRRLFVAELCSWMLMTFTLFTYDFVGEGL 300
DB 241 cpcrarlafnlgallprlhqccrmptlrrlrfvaeelcswmlmtlftlftdyvgegl 300
QY 301 YQGVRAEPGTEARRHYDEGRMGSLGLFLQCAISLVESLVMDRLVOREGTRAVYLASVA 360
DB 301 yqgvraepgtearrhydegvmgsllglflqcaislveslvmdrlyvregtravylasva 360
QY 361 AFPVAGATCISHSVAAVVTASAAALTGFTFSALQIIPYTLASLYHREKQVFLPKYRGDTGG 420
DB 361 afpvaagatcishsvaavvtasaaaltgftfsalqilpytlaalyhrekqvflpkrygdtgg 420
QY 421 ASSEDSLMTSFLPGPKPGAPPNNGHVAGSGSLPPPALCGASCDVSVRVVGEPTPEA 480
DB 421 assedslmtsfllpgpkpgapfnnghvagsgslppppalcgascdsvrvvgeptpea 480
QY 481 RVVPGRGICLDLALIDSAPFLSQVAPSLFMGSIYVLSOSVTAVMVSAAGLIVATYFAAQ 540
DB 481 rvvpgrgicldalidsafllsqvapslfmgsivqlsgsvtcaymvsaaglgivatyfatq 540

OY 541 VFEDKSDIAKYSA 553
 |||||||
 DB 541 vvfcdksdiakysa 553

RESULT 3
 AAB28527
 ID AAB28527 standard; Protein: 553 AA.

AC AAB28527;

DT 07-FEB-2001 (first entry)

DE Protein encoded by human breast tumour cDNA clone P501S.

KM Human; breast tumour antigen; cytostatic; immunotherapy;
 KW breast cancer; vaccine.

OS Homo sapiens.

PN WO200061756-A2.

PD 19-OCT-2000.

PF 10-APR-2000; 2000WO-US09688.

PR 09-APR-1999; 99US-0288950.

PR 02-JUL-1999; 99US-0346327.

PA (CORI-) CORIXA CORP.

PI Reed SG, Xu J, Dillon DC;

DR WPI: 2000-638568/61.

DR N-PSDB; AAC79473.

PT A novel isolated polypeptide comprising an immunogenic portion of a
 PT breast cancer protein useful in the detection and treatment of breast
 PT cancer -

PS Claim 2; Page 92-93; 95pp; English.

CC The present sequence is encoded by a cDNA sequence which was isolated
 CC from a breast tumour cDNA library. It is provided in a specification
 CC relating to compounds for immunotherapy and diagnosis of breast cancer.
 CC Breast tumour antigens and the polynucleotides that encode them may be
 CC used in the production of a pharmaceutical composition to be used in the
 CC treatment of breast cancer. Proliferated T cells and incubated antigen
 CC presenting cells are also required. The polypeptides and polynucleotides
 CC may also be used to produce a vaccine.

XX Sequence 553 AA;

SO

Query Match 100.0%; Score 2861; DB 21; Length 553;
 Best Local Similarity 100.0%; Pred. No. 1.8e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MVQRLVSRLLRRRAQDLVNLTFGLVCLAGITVPPPLLEVGVEKFMVLTIG 60
 |||||||
 DB 1 mvqrlvwarllrhrkaqlllvtllfglevclaagltvpplllevgyeekfmltvlg 60
 |||||||
 OY 61 PVLGIVCPPLGASADHMRGRYGRRRPTWALSIGLISLFLIPRAGWLAGLCPDPR 120
 |||||||
 DB 61 pvlglvcppllgasadhmrgrgrrrptfwalslgllslflfpragwlaqlcpdpr 120
 |||||||
 OY 121 ELALLILVGLDFCGQVCFPLLEALSDFRDPDHCROAVSVAFMISIGCLGYLPA 180
 |||||||
 DB 121 elallilvgldfcgqvcfpllealsdlfrdpdhrgrqavsvafmislgyclgyllpa 180
 |||||||
 OY 181 IDWTSALAPYIGTQECFLGILLIFLTCVATLLVAEALGTEPAEGISAPLSLPH 240
 |||||||

DB 181 idwtsalapyigtgeecflgilltlfltcvaatlivaeeaalgppeaglsapslsph 240
 |||||||
 OY 241 CCPCRARLAFRNIGALLPRLHQLCCMRPTLRFLFAELCSWMAALMTFTTDEYGECL 300
 |||||||
 DB 241 ccpccrariafrnlgallprlhqccmrptlrrlflvaelswmaalmftfltydftvg 300
 |||||||
 OY 301 YQGVPAEPGTFAARRHYDGVNRGSLGLFLOCAISLVSFLVMDRLVQRGTRAVYLA 360
 |||||||
 DB 301 yqgvpaepgtearrhydegvnmrgslglflocalslvsflvmdrlvqrgtravyla 360
 |||||||
 OY 361 APFVAAGATCLSHVAVVAWTAALGFTFSALQILPYTLASLYHREKQVFLPKRYGDTG 420
 |||||||
 DB 361 afpvaagatclshsvavvtaaalgtftfsalqilpytlaslhyrekyflpkrygdtg 420
 |||||||
 OY 421 ASSEDSIMTSEFLPGKPPGAPFNGHVGAGSGSLPPPALCGASACDVSVRVVGEPTA 480
 |||||||
 DB 421 asseedsimtsflpgkppgapfngvhgagsgslpppalcgasacdvsvrvvgepta 480
 |||||||
 OY 481 RVPVGRGICLDLAIDSAFLISOVAPSLFMGSIYOLSOQSVTAYMYSAAGLGVATYFATQ 540
 |||||||
 DB 481 rvppvgrgicldlaidsaflisovapslfmgsiyolsosqsvtaymysaaglgvatyfatq 540
 |||||||
 OY 541 VFEDKSDIAKYSA 553
 |||||||
 DB 541 vvfcdksdiakysa 553

RESULT 4
 AAY82002
 ID AAY82002 standard; Protein: 553 AA.

AC AAY82002;

DT 13-JUN-2000 (first entry)

DE Human immunogenic prostate tumour protein sequence SEQ ID NO:113.

KM Human; prostate cancer; diagnosis; tumour; gene therapy; detection;
 KW immunogenic; cytostatic; vaccine.

OS Homo sapiens.

PN WO200004149-A2.

PD 27-JAN-2000.

PF 14-JUL-1999; 99WO-US15838.

PR 14-JUL-1998; 98US-0115453.

PR 14-JUL-1998; 98US-0116134.

PR 23-SEP-1998; 98US-0158812.

PR 23-SEP-1998; 98US-0159822.

PR 15-JAN-1999; 99US-0232149.

PR 15-JAN-1999; 99US-0232880.

PR 09-APR-1999; 99US-0288946.

PA (CORI-) CORIXA CORP.

PI Dillon DC, Harlocker SL, Yung J, Xu J, Mitcham JL;

DR WPI: 2000-171268/15.

PT New polypeptide useful for treating and diagnosing prostate cancer
 PT comprises an immunogenic portion of prostate tumor protein -

CC Claim 3; Page 138-139; 263pp; English.
 CC The present invention describes isolated polypeptides, comprising an
 CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
 CC and polynucleotides encoding them have cytostatic activity and can be
 CC used in vaccines and in gene therapy. The polypeptides and
 CC polynucleotides encoding them, antigen presenting cells which express
 CC the polypeptides, antibodies against the polypeptides and vaccines

comprising them can be used for inhibiting the development of prostate cancer in a patient. The polypeptides can be used to generate antibodies or anti-idiotypic antibodies for passive immuno therapy. A portion of the polynucleotides encoding the polypeptides can be used as a probe or to modulate the expression of the polypeptides. AA06241 to AA06691 and CC AA082000 to AA082020 represent sequences used in the exemplification of the present invention.

Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 21; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.8e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MWQRIMWSRLRHRKAOULLVNLTFGLEVCLAGITVPPLLEVEGEKFMVIGIG 60
DB 1 mvgrlwsvrllrhkqallvnlftglveclaaqilyvpplllevegeekfmmvlgig 60
QY 61 PVLGLVCVPLIGSASDHMRGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 pvlglvcvplllgssasdhmrgryrgrrrpfifwalsigllslflipragwlagllcpdprpl 120
QY 121 ELALLILGVGLDFCGGYCTPLEALLSDLFRRDPDHCROAVSYAFMISLGGCGYLLPA 180
DB 121 elallilgvglldfcggyctfpleallsdldrddpchrqavsyafmislggcgyllpa 180
QY 181 IDMTSALAPYLGQECLEGLFLFTCVAAATLLVAEEALGPTPEAGLSAPSISPH 240
DB 181 idwtsalapylgqecclfgllftlftcvaatllvaeaalptepaeglsapsisph 240
QY 241 CCPCRARLAFRNLAGLRLHQLCCRMPTLRRLFVAELCSWMAIMTFTLYTDFVGEGL 300
DB 241 ccpcrarlafnrlagallprlhqlccmptlrrlrfvaeelcswmaimftlftdyfvegl 300
QY 301 YGVPRAPGTEARRHYDEGVKMSGLFLQCAISLVFSLYMDRLVQRFGRAYVLASVA 360
DB 301 ygvprapgtearrhydegvrmkslglflqcaislvsfslvmdrlvqrfgrayvlaava 360
QY 361 AEPFAAGATCISHVAVVTSAAALTGTFESALQILPYTIALSVREKQVFLPKRGDTGG 420
DB 361 afpfaagatcishvavvtasaaaltgtfesaqlilpytiaslvrekvflpkrgdtgg 420
QY 421 ASSDSLMTSFLPGPKGAPFPNGHVAGSGSGLLPPPALCGASACDVSVRVVVEPTEA 480
DB 421 asdsdsltstflpgpkgapfpnghvagsgslppppalcgasacdvsvrvvveptea 480
QY 481 RRVVGRGICDLALILDSAFLLSQVAPSLFMSIYQLSQSVTAHVWSAGLGLVAFATQ 540
DB 481 rrvvgrgicdlalildsafllsqvapslfmsiyqlsqsvtahvwsaaglgilvafatq 540
QY 541 VVFDKSDLAKEYSA 553
DB 541 vvfdksdlakysa 553

```

RESULT 5

AAU69763 standard; Protein: 553 AA.

AAU69763:

30-JAN-2002 (first entry)

Human prostate cDNA encoded protein #3.

Human: prostate cancer; cytosolic; immunostimulant; tumour; immunogen.

Homo sapiens.

MO200173032-A2.

04-OCT-2001.

27-MAR-2001; 2001MO-US09919.
27-MAR-2000; 2000US-0536857.
09-MAY-2000; 2000US-0568100.
12-MAY-2000; 2000US-0570737.
13-JUN-2000; 2000US-0593793.
27-JUN-2000; 2000US-0605783.
10-AUG-2000; 2000US-0636215.
29-AUG-2000; 2000US-0651236.
06-SEP-2000; 2000US-0657279.
02-OCT-2000; 2000US-0679426.
10-OCT-2000; 2000US-0685166.
(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD; Finger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D; Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA; WPI; 2001-639232/73.
N-PDB; AAS63557.

New human prostate-specific polypeptides and polynucleotides useful for the diagnosis and treatment of cancer, especially prostate cancer - Claim 2; Page 269-270; 579pp; English.

The invention relates to isolated prostate-specific polynucleotides, polypeptides, fusion proteins of the polypeptides, antibodies raised against the polypeptides (or antigenic epitopes derived from them) and antigen-presenting cells expressing the polypeptides. The antibodies are useful for detecting the presence of cancer, especially prostate cancer. The polypeptides, polynucleotides and the antigen-presenting cells are useful for stimulating and/or expanding T cells specific for a tumour protein, and for inhibiting the development of cancer especially prostate cancer. Compositions comprising the polynucleotide and/or polypeptide are useful for stimulating an immune response, and for treating cancer. The oligonucleotide is useful for detecting cancer. The present sequence is a prostate specific polypeptide of the invention.

Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.8e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MWQRIMWSRLRHRKAOULLVNLTFGLEVCLAGITVPPLLEVEGEKFMVIGIG 60
DB 1 mvgrlwsvrllrhkqallvnlftglveclaaqilyvpplllevegeekfmmvlgig 60
QY 61 PVLGLVCVPLIGSASDHMRGRYGRRRPFIWALSIGILLSLFLIPRAGWLAGLCPDPRPL 120
DB 61 pvlglvcvplllgssasdhmrgryrgrrrpfifwalsigllslflipragwlagllcpdprpl 120
QY 121 ELALLILGVGLDFCGGYCTPLEALLSDLFRRDPDHCROAVSYAFMISLGGCGYLLPA 180
DB 121 elallilgvglldfcggyctfpleallsdldrddpchrqavsyafmislggcgyllpa 180
QY 181 IDMTSALAPYLGQECLEGLFLFTCVAAATLLVAEEALGPTPEAGLSAPSISPH 240
DB 181 idwtsalapylgqecclfgllftlftcvaatllvaeaalptepaeglsapsisph 240
QY 241 CCPCRARLAFRNLAGLRLHQLCCRMPTLRRLFVAELCSWMAIMTFTLYTDFVGEGL 300
DB 241 ccpcrarlafnrlagallprlhqlccmptlrrlrfvaeelcswmaimftlftdyfvegl 300
QY 301 YGVPRAPGTEARRHYDEGVKMSGLFLQCAISLVFSLYMDRLVQRFGRAYVLASVA 360
DB 301 ygvprapgtearrhydegvrmkslglflqcaislvsfslvmdrlvqrfgrayvlaava 360

```

OY 361 AFPVAGATCLSHVAVVTASAAITGFTFSAIOLITPTLASLYHREKQVFLPKYRGDTG 420
 DB 361 afpvaagatclshvavvtasaaITGFTFSAIOLITPTLASLYHREKQVFLPKYRGDTG 420
 OY 421 ASSEDSIMTSFLPGPKGAPFPNGHVAGSGSLPPPALCGASACDVSVRVVVGPEPTA 480
 DB 421 asseDSIMTSflpgpkgapfpnghvagsgslppppalcgasacdvsvrvvvgeptea 480
 OY 481 RVVPGRGICLDLALIDSAPFLSQVAPSLFMGSIYOLSOVTAIYVSAAGLVAIYFATQ 540
 DB 481 rvpgrgicldlaldsAFLSQVAPSLFMGSIYOLSOVTAIYVSAAGLVAIYFATQ 540
 OY 541 VVFDKSDIAKYSA 553
 DB 541 vvfdksdiakysa 553

RESULT 6

AAU04961
 ID AAU04961 standard; Protein: 553 AA.

AC AAU04961;
 DT 24-OCT-2001 (first entry)
 DE Human prostate tumour protein L1-12.
 KW Human: prostate tumour protein; prostate cancer.
 OS Homo sapiens.
 PN US6262245-B1.
 PD 17-JUL-2001.
 PF 25-FEB-1998; 98US-0030607.
 PR 25-FEB-1997; 97US-0806099.
 PR 01-AUG-1997; 97US-0904804.
 PR 09-FEB-1998; 98US-0020956.
 PA (CORI-) CORIXA CORP.
 PI Xu J, Dillon DC;
 DR N-PSDB; AAS10108.
 DR WPI: 2001-440862/47.
 PT Novel polynucleotide encoding polypeptide comprising a portion of prostate tumour protein useful for inhibiting development of prostate cancer or for treating prostate cancer in a patient
 PS Example 1; Column 125-127; 105pp; English.
 CC The sequence is a partial prostate tumour protein, encoded by a prostate tumour specific cDNA. The DNA is useful for inhibiting the development of prostate cancer or for treating prostate cancer in a patient.
 CC of prostate cancer or for treating prostate cancer in a patient.
 SQ Sequence 553 AA;

Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 1.8e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MOWRLVSRLLRRKAKQLLNLITFGLVCLAGITVYPLLVEGVEKFMVWLGIG 60
 DB 1 mowrlvsrllrrrkakqlnlitfglvclagitvyppllvegyvekfmvwlgiq 60
 OY 61 PVIGLVCVPLIGSASDHWRGRRRFTWALSGLISLFLIPRAGWLAGILCPDPRPL 120
 DB 61 pviGLVCVPLIGSASDHWRGRRRFTWALSGLISLflIPRAGWlaGIlCPDpRpL 120

OY 121 ELALLITGVGLIDFGQVCFPLEALLSDLFDPDPHCRQAVSVAFMISLGCCIGYLLPA 180
 DB 121 elallitgvglidfgqvctfpleallsdlfddpdpchrqavsvafmislGCClgyllpa 180
 OY 181 IDWDTSALAPVLTQRECEFLGTLTLFLNCVAATLVAEBALGPEPAEGISASLSPH 240
 DB 181 idwDtsAlApVltqReceflgtlTLflNCvAATLVAEBALGPEPAEGISASlSPH 240
 OY 241 CQPCRARLARFNLGALLPRHOLCCMPTIRLFAELCSMMALMTFTLFYTDVEGEG 300
 DB 241 cqpCraRLaRfNLgAlLpRhOLccMptIRlFAELCSMMAlMTFTlFYTDVEgeG 300
 OY 301 YQGVRAEPGTEARRHNDGVRMGSGLFLQCAISLVFSIMDRLOVORFTRAVIYASVA 360
 DB 301 ygvraepgtearrhndgvrmgsglflqcaislvsfSimDRLOVORfTRaVIYASva 360
 OY 361 AFPVAGATCLSHVAVVTASAAITGFTFSAIOLITPTLASLYHREKQVFLPKYRGDTG 420
 DB 361 afpvaagatclshvavvtasaaITGFTFSAIOLITPTLASLYHREKQVFLPKYRGDTG 420
 OY 421 ASSEDSIMTSFLPGPKGAPFPNGHVAGSGSLPPPALCGASACDVSVRVVVGPEPTA 480
 DB 421 asseDSIMTSflpgpkgapfpnghvagsgslppppalcgasacdvsvrvvvgeptea 480
 OY 481 RVVPGRGICLDLALIDSAPFLSQVAPSLFMGSIYOLSOVTAIYVSAAGLVAIYFATQ 540
 DB 481 rvpgrgicldlaldsAFLSQVAPSLFMGSIYOLSOVTAIYVSAAGLVAIYFATQ 540
 OY 541 VVFDKSDIAKYSA 553
 DB 541 vvfdksdiakysa 553

RESULT 7

AAU01117
 ID AAU01117 standard; Protein: 553 AA.

AC AAU01117;
 DT 04-OCT-2001 (first entry)
 DE Human prostate-specific amino acid sequence L1-12.
 KW Human: prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis.
 OS Homo sapiens.
 PN WO200151633-A2.
 PD 19-JUL-2001.
 PF 16-JAN-2001; 2001WO-US01574.
 PR 14-JAN-2000; 2000US-0483672.
 PA (CORI-) CORIXA CORP.
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SC, Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW, Wang A, Meagher MD;
 DR WPI: 2001-425873/45.
 PT New polynucleotide encoding a prostate-specific protein, for diagnosis, monitoring and treating prostate cancer in a patient and for use in vaccines
 PS Claim 2; Page 267-268; 543pp; English.
 CC The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II),

CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
CC AAH01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.

XX Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.8e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRRKQQLLVNLTFFGEVCLAAGITVYPPILLEVEGEKEMVNLGIG 60
DB 1 mvgrlwvsrllrrkqqlvnlftfglevclaagityvppilllevgeekfmlvlgig 60
QY 61 PVLGLVCPVLGSASDHWGRYGRRRPPTWALSIGILSLFIPRAGWLAGLCPDPRPL 120
DB 61 pvlglvcvplgsasdhwgrgygrrrpftwalsigilslfipragwlagllcpdprpl 120
QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPHCRQAVSVAFMISLGGCGLYLPA 180
DB 121 elallilgvglldfcgvcfpleallsdldpdpchrcqavsvafmislggcglylpa 180
QY 181 IMWDTSAAPYVIGTQECFLGLTLFLTCVAATLLVAEALGPTPEAGLSAPSISPH 240
DB 181 idwdsalapyigtqecflgltlfltcvaatllvaealgp tpeaglsapsisph 240
QY 241 CCPCARLAFRNALGALPRLHQLCCMRPTRLRFLVAFELCSWMAIMFTFLFYTFVGGEL 300
DB 241 ccpcarlatfnlga llprlhqllccmrptrllrfvaelcswmalmtflfytfvgeg l 300
QY 301 YQGVRAEPGTGARHDEGRVMSLGLFLQCAISLVFSVMRLVORFTRAVYLA SVA 360
DB 301 ygvpraepgtearhdegvrmslglflqcaisl vfstvmrldvorftravylasva 360
QY 361 AFPVAAGATCISHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKRGDTGG 420
DB 361 afpvaagatcishsvavvtasaltgftfsalqilpytlaslyhrekqvflpyrgdtgg 420
QY 421 ASSDSILMTSFLPRKRGAPPPNGHVAGSGGLPPPALCGASACDVSVRVVGEPTEA 480
DB 421 assdsilmtsf lprkrgapppnghvagsgglpppalcgasacdvsvrvvgeptea 480
QY 481 RVPBGRICLDLALIDSAPFLSQVAPSLFMGSTVOLSOSTAYVWSAAGLVAIFYATQ 540
DB 481 rvpbgricldlaldsapflsqvapsl fmgstvolstoyvwsaaglvaiyfataq 540
QY 541 VVPEKSLAYKSA 553
DB 541 vvpekslaky sa 553

RESULT 8
ID AAG99002
XX AAG99002 standard; Protein; 553 AA.
AC AAG99002;

25-SEP-2001 (first entry)

prostate-specific amino acid sequence L1-12/p501S.

prostate cancer; therapy; diagnosis; cat eye syndrome;
L1-12; prostate-specific protein; chromosome 1;
antigen; PSA.

XX Homo sapiens.
OS
XX
XX W0200134802-A2.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000WO-US30904.
XX
XX 12-NOV-1999; 99US-0439313.
XX
XX 18-NOV-1999; 99US-0443686.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG,
XX Katos MD, Retter MW, Stolk JA, Day CH, Skelky YAW, Wang A;
XX WPI; 2001-308785/32.

Isolated polypeptide comprising at least an immunogenic portion of a
prostate-specific protein, useful in the diagnosis and therapy of
prostate cancer -

Claim 3; Page 167-168; 325pp; English.

The present invention describes an isolated polypeptide (P1) comprising
at least an immunogenic portion of a prostate-specific protein, or its
variant. Also described are polynucleotides (N1) encoding (P1). (P1) and
(N1) have cytostatic activity and can be used in vaccine production.
The polypeptides, nucleic acids and antibodies from the present
invention are useful in the diagnosis and therapy of prostate cancer.
Prostate specific genes P704P, P712P, P774P, P775P and B305P are located
in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
region. Prostate specific antigen (PSA) p501S was located on
chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
polynucleotide and polypeptide sequences used in the exemplification
of the present invention.

XX Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 1.8e-269;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MVORLWVSRLLRRKQQLLVNLTFFGEVCLAAGITVYPPILLEVEGEKEMVNLGIG 60
DB 1 mvgrlwvsrllrrkqqlvnlftfglevclaagityvppilllevgeekfmlvlgig 60
QY 61 PVLGLVCPVLGSASDHWGRYGRRRPPTWALSIGILSLFIPRAGWLAGLCPDPRPL 120
DB 61 pvlglvcvplgsasdhwgrgygrrrpftwalsigilslfipragwlagllcpdprpl 120
QY 121 ELALLILGVGLDFCGQVCFPLEALLSDLPDPHCRQAVSVAFMISLGGCGLYLPA 180
DB 121 elallilgvglldfcgvcfpleallsdldpdpchrcqavsvafmislggcglylpa 180
QY 181 IMWDTSAAPYVIGTQECFLGLTLFLTCVAATLLVAEALGPTPEAGLSAPSISPH 240
DB 181 idwdsalapyigtqecflgltlfltcvaatllvaealgp tpeaglsapsisph 240
QY 241 CCPCARLAFRNALGALPRLHQLCCMRPTRLRFLVAFELCSWMAIMFTFLFYTFVGGEL 300
DB 241 ccpcarlatfnlga llprlhqllccmrptrllrfvaelcswmalmtflfytfvgeg l 300
QY 301 YQGVRAEPGTGARHDEGRVMSLGLFLQCAISLVFSVMRLVORFTRAVYLA SVA 360
DB 301 ygvpraepgtearhdegvrmslglflqcaisl vfstvmrldvorftravylasva 360
QY 361 AFPVAAGATCISHSVAVVTASALTGFTFSALQILPYTLASLYHREKQVFLPKRGDTGG 420
DB 361 afpvaagatcishsvavvtasaltgftfsalqilpytlaslyhrekqvflpyrgdtgg 420

OY 421 ASSEDSIMTSFLPGPKGAPFPNGHVAGAGSGLLPPPALCGASACDVSVRVVGEPTA 480
 DB 421 asseedsimtsflpgpkpgapfpngvhvgaagsgllppppalcgasacdvsvrvvgeptea 480
 OY 481 RVPVGRGICLDLAIDSAFLSQVAPSLFMGSIYQLSGSVTAYMWSAAGLGLVAIYFATQ 540
 DB 481 rvpvgrgicldlaidsaflsqvapslfmgsivqsgsvtaymwsaaglgivaifacq 540
 OY 541 VVFEDKSDLAKYSA 553
 DB 541 vvfedksdlakysa 553

RESULT 9

AAG62150
 ID AAG62150 standard: Protein; 553 AA.

XX AAG62150;

DT 06-JUL-2001 (first entry)

DE Human P501S Inventive antigen SEQ ID NO: 333.

XX Human; mouse; immunotherapy; cancer; leukaemia; WTI; wilm's tumour gene;
 KW chromosome 11p13; zinc finger transcription factor.

OS Homo sapiens.

WO200125273-A2.

PD 12-APR-2001.

PF 04-OCT-2000; 2000MO-US27465.

PR 04-OCT-1999; 990S-0157459.

PA (CORI-) CORIXA CORP.

PI Skelky YAW, Xu J, Cheever MA, Reed SG;

DR WPI; 2001-328324/34.

PT Polypeptide comprising part of the Wilm's Tumour gene product sequence is
 used in the diagnosis and treatment of malignant diseases e.g. leukemia
 and cancer associated with WTI -

PS Disclosure: Page 212-213; 228pp; English.

XX The present invention describes compositions comprising peptides derived
 CC from the Wilm's tumour protein WTI and methods for their use in treating
 CC malignant diseases. Peptides derived from both the murine and human WTI
 CC proteins are provided. The human WTI gene is found on chromosome 11p13,
 CC and the protein was shown to be a zinc finger transcription factor. The
 CC immunogenic peptides of the invention are particularly useful in the
 CC diagnosis and treatment of cancer and leukaemia. The present sequence is
 CC a polypeptide described in the exemplification of the invention.

XX Sequence 553 AA:

Query Match 100.0%; Score 2861; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 1.8e-269;
 Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MYQRLWVSRILRRKQOLLVNLITFGLVYCLAAGITVYPPPLLEVEKEMTMYLGIG 60
 DB 1 myqrlwvsrllrrkqollvnlitfglevclaagitypppllleveekemtmylgig 60

OY 61 PVLGLVCPVLIGSASDHMGVGRGRBPETWALSIGLISFLTPRAGMLAGLCPDPRPL 120
 DB 61 pvlglvcvpllgsasdhmgvgrgrbpetwalsigllsfltpragmlaglcppdprpl 120

OY 121 ELALLILGVGLDFCGVCFPTPLALLSDLFRPDHCRQAVSYVAFMISLGGCLGYLLPA 180

DB 121 elallilgylldfcgvcftpleallsdldfrpdhcrqavsyvatfmlsiggclgyllpa 180
 OY 181 IMDPISALAPYLGTOEHCIFGLITLFLTCVAATLLVAEAAIGPREPREGISAPSLSPH 240
 DB 181 imdpsalapylgtoehcifglitlfltcvaatllvaeaaigpreperegisapslsp 240
 OY 241 CCECRARLAFRNLGALPLPHOLCCRMPTLRLEFVAELCSMMALMTFTLTDFEYGEGL 300
 DB 241 ccecrarlafnrnlgalplpholccrmptlrllefvaelcsmmalmtftltdfeygegl 300
 OY 301 YQGVPRAEDETEARRHYDEGVNMGSLGLFLQCAISLVSFLVMDRLVQPRGTRAVIYASVA 360
 DB 301 yqgvpraedetearrhydegvnmgsiglfqcaislvsflvmdrlvqprgtraviasva 360
 OY 361 APPVAGATCLSHSVAVVYASALNGETPSALQIIPYTLASLYHREKQVFLPKYREDTNG 420
 DB 361 appvagatclshsvavvyasalngetpsalqipyltlaslyhrekqvfllpkyyrdtng 420
 OY 421 ASSEDSIMTSFLPGPKGAPFPNGHVAGAGSGLLPPPALCGASACDVSVRVVGEPTA 480
 DB 421 asseedsimtsflpgpkpgapfpngvhvgaagsgllppppalcgasacdvsvrvvgeptea 480
 OY 481 RVPVGRGICLDLAIDSAFLSQVAPSLFMGSIYQLSGSVTAYMWSAAGLGLVAIYFATQ 540
 DB 481 rvpvgrgicldlaidsaflsqvapslfmgsivqsgsvtaymwsaaglgivaifacq 540
 OY 541 VVFEDKSDLAKYSA 553
 DB 541 vvfedksdlakysa 553

RESULT 10

AAB74800
 ID AAB74800 standard: Protein; 553 AA.

XX AAB74800;

DT 14-JUN-2001 (first entry)

DE Prostate tumour antigen predicted amino acid sequence for LI-12.

XX Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
 KW prostate cancer; immunogenic; cytostatic; vaccine.

OS Homo sapiens.

WO200125272-A2.

PD 12-APR-2001.

PF 04-OCT-2000; 2000MO-US27464.

PR 04-OCT-1999; 990S-0157455.

PA (CORI-) CORIXA CORP.

PI Xu J, Skelky YAW, Reed SG, Cheever MA;

DR WPI; 2001-245062/25.

DR N-PSDB; AAH02530.

PT Prostate specific protein and its encoding polynucleotide, useful for
 the treatment and diagnosis of prostate cancer -

PS Claim 3; Page 157-158; 276pp; English.

XX The present invention describes an isolated polypeptide (I) comprising
 CC at least an immunogenic portion of a prostate tumour antigen protein or
 CC its variant. (I) have cytostatic activity and can be used in vaccine
 CC production. (I), prostate tumour antigen polynucleotides, an antigen
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
 CC pharmaceutical composition containing (I) are useful for inhibiting the

OY 541 VVFDKSDIAKSYA 553
Db 1067 vvfksdlsakysa 1079

RESULT 12

AA001318 standard; Protein: 595 AA.

AA001318:

04-OCT-2001 (first entry)

Alpha prepro-P501S recombinant protein amino acid sequence.

Human; prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis.

Homo sapiens.

WO200151633-A2.

19-JUL-2001.

16-JAN-2001; 2001WO-US01574.

14-JAN-2000; 2000US-0483672.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG, Katos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW, Wang A, Meagher MJ;

WPI; 2001-425873/45.

New polynucleotide encoding a prostate-specific protein, for

diagnosing, monitoring and treating prostate cancer in a patient and

for use in vaccines -

Example 17; Page 541-543; 543pp; English.

The present invention describes polynucleotide sequences (I) which encode prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production and gene therapy. (I), (II), antibodies to (I), fusion proteins comprising (I), and isolated T cells prepared using (I) or (II) are used treat cancer in a patient. (I) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is particularly prostate cancer. (I) and (II) can be used in vaccines. The antibodies or (I) can be used for monitoring the progression of cancer in a patient. (I) and (II) can also be used to improve diagnostic and therapeutic methods for prostate cancer. They can indicate the level of metastasis as well as the prostate volume. AA093357 to AA093944 and AA001115 to AA001318 represent polynucleotide and amino acid sequences used in the exemplification of the present invention.

Sequence 595 AA;

Query Match 90.7%; Score 2596; DB 22; Length 595;
Best Local Similarity 95.1%; Pred. No. 1.2e-243;
Matches 501; Conservative 9; Mismatches 9; Indels 8; Gaps 1;

OY 35 GIFYVPPLLLEVGEKEMT-----WVIGIPVGLVGVPLLGSSSDHRRGKRYGRR 86
Db 61 gllfntltlaakkeagyslekreaamvlgivpvlglvcpvllgsasdhwtgrygrr 120
OY 87 PRWALSGLTLLSLFIPRAGWLAGLCPDPRFLALLLGVGLDFCGVCFTRPLAL 146
Db 121 pftwalslgllstflfipragwlaglcpdprlatallllgvglldfegvcfcftrplaal 180

OY 147 LSDLFRDPDCHROAYSVAAPMISLGGCLGYLPAIDMTSALAPYLGTOECLFGLTLI 206
Db 181 lsdlfrdpdchrcqaysvayamlsiggclgyllpaiddcsalapylgteoclfglltli 240
OY 207 FLTCVANTLVAEEMALGPTEPAEGLSAPSLSPHCCPCRARLAFRLGALIPRLHQLCR 266
Db 241 fltcvaantllvaeemalgppepaeglsapslsphccpcrarlafrrlgaalprlhaqlcr 300
OY 267 MPRTLRLFLVAELCSKMAAMFTLFTDFYGEGLYGVPAEPTEARRRHHYDEGVHMSL 326
Db 301 mprtlrlflvaelcskmaamftlftdfygegllygvpaepptearrrhhydegvmgs 360
OY 327 GLFLQCAISLVESLVMDLVOREGTRAVYLASVAAPVAAGATCLSHSAVVTASALTLG 386
Db 361 glflqcaislveslvmdlvoregtravylasvaapvaagatclshsavvtasaltlg 420
OY 387 FTFSALQILPYTLASLYHREKQVFLPKYRQDTGASSEDLSMTSLFPGKAPFNGHY 446
Db 421 ftfsalqlpytlaslyhrekqvlpkryrdtggasedsimtsflpgkappfngyh 480
OY 447 GAGSGILPPPALCGASACVSVRVVYGEPTEARVYVPGRGICLDLAILDSAEFLSQVAP 506
Db 481 gagsgilpppalcgasacsvrvvvgeptearvypvgrgicldlaildsafllsqvap 540
OY 507 SLFMGSIVQLSQSVYAVVWASAGLVAIYFATQVDFDKSDIAKSYA 553
Db 541 slfmgshivqlsqsvyavvwasaglvaiyfataqvdfksdlsakysa 587

RESULT 13

AA069875 standard; Protein: 371 AA.

AA069875:

30-JAN-2002 (first entry)

Human prostate cDNA encoded protein #69.

Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.

Homo sapiens.

WO200173032-A2.

04-OCT-2001.

27-MAR-2001; 2001WO-US09919.

27-MAR-2000; 2000US-0536857.
PR 09-MAY-2000; 2000US-0568100.
PR 12-MAY-2000; 2000US-0570737.
PR 13-JUN-2000; 2000US-0593793.
PR 27-JUN-2000; 2000US-0605783.
PR 10-AUG-2000; 2000US-0636215.
PR 29-AUG-2000; 2000US-0651236.
PR 06-SEP-2000; 2000US-0657279.
PR 02-OCT-2000; 2000US-0679426.
PR 10-OCT-2000; 2000US-0685166.

(CORI-) CORIXA CORP.

Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Katos MD; Fanger GR, Retter MW, Stolk JA, Day CH, Vedrick TS, Carter D; Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;

WPI; 2001-639232/73.
N-PDSB; AAS64038.

New human prostate-specific polypeptides and polynucleotides useful for the diagnosis and treatment of cancer, especially prostate cancer -
Claim 2; Page 487-488; 579pp; English.

50 Sequence 371 AA;

| | | | | | | | | | |
|---------|------|--------------|----|------------|-----|--------|-----|------|---|
| Matches | 269; | Conservative | 6; | Mismatches | 11; | Indels | 13; | Gaps | 1 |
|---------|------|--------------|----|------------|-----|--------|-----|------|---|

Db 9 slvplplalsqplthtsll-----aglgpvlglvcpplqsasdhwrqr 55

50 ygrirprlwalslgllslflpragwlagllcpdprrleal11lgvg1ldfcgvcft 115

175

-----accuuypceaylsapsispncpcrarlalnlgallprln 235

294

| ID | Protein | AA |
|----------|-------------------|---------|
| AAM01230 | standard; Protein | 371 AA. |

DI 04-OCT-2001 (first entry)
XX

human; prostate cancer; prostate-specific; diagnosis; vaccine; cytostatic; gene therapy; metastasis

PN WO200151633-A2.

PF 16-JAN-2001; 2001WO-US01574.

(CORI-) CORIXA CORP.

Meagher MJ;

673/45.

PS Claim 2; Page 464-466; 543pp; English.

prostate-specific proteins (II). (I) and (II) have cytostatic activity, and can be used in vaccine production.

(1) and the antibodies are also used in the detection of cancer in a patient. The cancer that is diagnosed or treated is substantially-

methods for prostate cancer. They can indicate the level of metastasis

NOT RECORDED

```
Query Match      49.58;  Score 1417.5;  DB 22;  Length 371;
PostgreSQL 9.4.1
```

22 NLLTFGLEVCLAGITYVPPLLLEVGEVEKEKMTMVLGIGPVLCVPLLSASDHWGR 81

02 1GNNFF1WALSDG1LUSLF1IPKAGWLAGLLCPDPRPLETALITGVGLDFCGVCFT 141

XX

[illegible][illegible]

RESULT: 15
AAE01362

AAE01382;
AC
XX

XXXXXX gene 14 encoded secreted protein HMBAK14, SEQ ID NO:84.

KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder:

pregnancy-related disorder; endocrine disorder; infection; wound healing;

OS Homo sapiens.
XX

Key Location/Qualifiers
 Peptide 1..48
 Protein /label= Signal_peptide
 Misc-difference 49..371
 /label= Mature_human_secreted_protein
 /label= Unknown
 /note= "Encoded by TNC"
 WO200134629-A1.
 17-MAY-2001.
 08-NOV-2000; 2000WO-US30654.
 12-NOV-1999; 99US-0164835.
 27-JUL-2000; 2000US-0221142.
 (HUMA-) HUMAN GENOME SCI INC.
 Ruben SM, Komatsoulis GA, Wei P, Fiscella M, Baker KP;
 MPI; 2001-308779/32.
 N-PSDB; AAD05230.
 New nucleic acid encoding one of 21 human secreted proteins for
 diagnosing, preventing, treating or ameliorating medical conditions,
 such as autoimmune disease and cancer, and used as a food additive or
 preservative -
 Claim 11; Page 438-440; 490pp; English.
 AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted
 protein genes, and AAE01352-AAE01413 represent the proteins they encode.
 AAE01415-AAE01433 represent human secreted protein fragments or variants.
 The secreted proteins and their genes are useful for preventing, treating
 or ameliorating medical conditions, e.g., by protein or gene therapy.
 Pathological conditions can be diagnosed by determining the amount of the
 new protein in a sample or by determining the presence of mutations in
 the new genes. Specific uses are described for each of the 21 genes,
 based on the tissues in which they are most highly expressed, and include
 developing products for the diagnosis or treatment of proliferative
 disorders, cancer, tumours, foetal and developmental abnormalities,
 haematopoietic disorders, diseases of the immune system, AIDS, autoimmune
 diseases (e.g., rheumatoid arthritis), inflammation allergies,
 neurological disorders (e.g., Alzheimer's disease, Parkinson's disease),
 cognitive disorders, schizophrenia, asthma, skin disorders (e.g.,
 psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders,
 angiotensin-related disorders, kidney disorders, gastrointestinal disorders,
 pregnancy-related disorders, endocrine disorders, and infections. The
 proteins can also be used to aid wound healing and epithelial cell
 proliferation, to prevent skin aging due to sunburn, to maintain organs
 before transplantation, for supporting cell culture of primary tissues,
 to regenerate tissues, to identify their cognate ligands or binding
 partners, and in chemotaxis, and can be used as a food additive or
 preservative to modify storage properties. Antibodies specific for a
 protein of the invention can be used in alleviating symptoms associated
 with the disorders mentioned above, and in diagnostic immunoassays e.g.,
 radioimmunoassay or enzyme linked immunosorbent assay (ELISA).
 The present sequence represents a human secreted protein of
 the invention.
 SQ Sequence 371 AA:
 Query Match 49.5%; Score 1416; DB 22; Length 371;
 Best Local Similarity 98.5%; Pred. No. 4,1e-129;
 Matches 263; Conservative 3; Mismatches 1; Indels 0; Gaps 0;
 OY 54 TMTLGGPVLGTCVPLSGASDHWGRGRGRRPFMTALSLGILSLFLIPRAGWLAGLL 113
 DB 28 sllagipvlgtvcplsgasdhwgrgrgrrpfmtalslgilslflipragwlagll 87

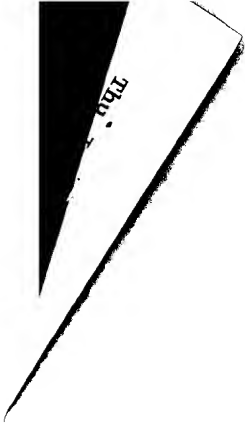
OY 114 CDPPELEALLITGVLDFGSGVCFTEPLEALLSDLFDRDPHCRQAVSVAFMISLGGC 173
 DB 88 cpdprrleallilgyglldfcgqvctfpleallslldfrdpdrcqavsvafmislsgc 147
 OY 174 LGYLPAIDMDTSALAPYLGTQECDFGLTLFLTCVAATLIVAEALGTPERAEGLS 233
 DB 148 lgyllpaidmdtsalapygtqecclfglltlfltcvaatlivaeaalgpereagls 207
 OY 234 APSLSPPCCCRARAFNRNIGALLPRLHOLCCMPRTLRLFFVAELCSMMALMTFTLYT 293
 DB 208 apslsphccpcrarlafnrnigallprlhqccmptlrlrlfvaelcsmmalmtftlft 267
 OY 294 DFVGEGLYGVRAEPCTEARRHYDEG 320
 DB 268 dfvgeglygvpraeptearrhydeg 294

Search completed: June 26, 2002, 13:43:01
 Job time: 4099 sec

Thu Jun 27 09:46:43 2002

us-09-838-785-2.rag

Page 12



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:37:17 ; Search time 87.43 Seconds
(without alignments)
1094.203 Million cell updates/sec

Title: US-09-838-785-2

Perfect score: 2861
Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVPEKSDIAKYS A 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_19:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_phage:*

10: sp_plant:*

11: sp_prodent:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_virus:*

16: sp_bacteriap:*

17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 2861 | 100.0 | 553 | 4 | Q96JT2 |
| 2 | 2798 | 97.8 | 553 | 6 | Q95K15 |
| 3 | 2531 | 88.5 | 501 | 6 | Q95K15 |
| 4 | 519.5 | 18.2 | 599 | 5 | Q9SVL1 |
| 5 | 456 | 15.9 | 576 | 13 | Q90Z74 |
| 6 | 347.5 | 12.1 | 515 | 10 | Q65803 |
| 7 | 347.5 | 12.1 | 515 | 10 | Q65803 |
| 8 | 346.5 | 12.1 | 501 | 10 | Q9FNR6 |
| 9 | 346.5 | 12.1 | 515 | 10 | Q9SOK6 |
| 10 | 346 | 12.1 | 612 | 10 | Q9SLN7 |
| 11 | 338 | 11.8 | 512 | 10 | Q9SOK5 |
| 12 | 337.5 | 11.8 | 523 | 10 | Q9SOK5 |
| 13 | 336.5 | 11.8 | 501 | 10 | Q9SP63 |
| 14 | 333 | 11.6 | 512 | 10 | Q9ZTB9 |
| 15 | 330 | 11.5 | 516 | 10 | Q43653 |
| 16 | 329.5 | 11.5 | 524 | 10 | Q9XHL6 |

| | | | | | |
|----|-------|------|-----|----|--------|
| 17 | 328.5 | 11.5 | 506 | 10 | Q944W2 |
| 18 | 328.5 | 11.5 | 594 | 10 | Q80605 |
| 19 | 327.5 | 11.4 | 506 | 10 | Q948L0 |
| 20 | 325.5 | 11.4 | 604 | 10 | Q9FVL6 |
| 21 | 323.5 | 11.3 | 512 | 10 | Q9J231 |
| 22 | 323.5 | 11.3 | 512 | 10 | Q80550 |
| 23 | 321.5 | 11.2 | 507 | 10 | Q9SP14 |
| 24 | 317.5 | 11.1 | 507 | 10 | Q40583 |
| 25 | 317.5 | 11.1 | 530 | 10 | Q9M535 |
| 26 | 316.5 | 11.1 | 500 | 10 | Q9FV06 |
| 27 | 313 | 10.9 | 521 | 10 | Q9SXM0 |
| 28 | 311 | 10.9 | 533 | 10 | Q41152 |
| 29 | 307.5 | 10.7 | 488 | 10 | Q9FV92 |
| 30 | 306 | 10.7 | 523 | 10 | Q9M422 |
| 31 | 305.5 | 10.7 | 501 | 10 | Q65929 |
| 32 | 302.5 | 10.6 | 510 | 10 | Q9SP15 |
| 33 | 301.5 | 10.5 | 429 | 10 | Q40167 |
| 34 | 299.5 | 10.5 | 505 | 10 | Q9SOK4 |
| 35 | 296.5 | 10.4 | 510 | 10 | Q9M3R4 |
| 36 | 294.5 | 10.3 | 510 | 10 | Q9FE59 |
| 37 | 294 | 10.3 | 510 | 10 | Q40938 |
| 38 | 294 | 10.3 | 523 | 10 | Q9J438 |
| 39 | 292.5 | 10.2 | 537 | 10 | Q9C8X2 |
| 40 | 292 | 10.2 | 537 | 10 | Q49838 |
| 41 | 292 | 10.2 | 538 | 10 | Q91KH3 |
| 42 | 291.5 | 10.2 | 509 | 10 | Q9FNR9 |
| 43 | 291 | 10.2 | 618 | 4 | Q9UL03 |
| 44 | 287 | 10.0 | 539 | 10 | Q9SM24 |
| 45 | 286 | 10.0 | 474 | 10 | Q04516 |

ALIGNMENTS

| | | | |
|--|---|------|---------|
| RESULT 1 | PRELIMINARY: | PRT: | 553 AA. |
| Q96JT2 | Q96JT2 | | |
| AC | Q96JT2: | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Created) | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last sequence update) | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) | | |
| DE | PROTEIN. | | |
| OS | Homo sapiens (Human). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. | | |
| OX | NCBI_TaxID=9606; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | TISSUE=PROSTATE; | | |
| RX | MEDLINE=21139094; PubMed=11245466; | | |
| RA | Xu J., Kalos M., Stoik J.A., Zasloff E.J., Zhang X., Houghton R.L., | | |
| RA | Filho A.M., Nolasco M., Badaro R., Reed S.G.; | | |
| RT | "Identification and characterization of protein, a novel prostate- | | |
| RT | specific protein." | | |
| RL | Cancer Res. 61:1563-1568(2001). | | |
| DR | EMBL, AY033593; AAK54386.1;.. | | |
| SQ | SEQUENCE 553 AA: 59322 MW: 0AFA23FBC742A667 CRC64: | | |
| Query Match | 100.0%: Score 2861; DB 4; Length 553; | | |
| Best Local Similarity | 100.0%: Pred. No. 5.4e-213; | | |
| Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | |
| QY | 1 MVQRLWVSRLLRHRKAQLLVNLTFFGLEVCIAAGITVPPPLLEVGVEEKEKMTVLGIG 60 | | |
| DB | 1 MVQRLWVSRLLRHRKAQLLVNLTFFGLEVCIAAGITVPPPLLEVGVEEKEKMTVLGIG 60 | | |
| QY | 61 PVLGLVCPPLGASADHMRGRGRRPRITWALSGLISLFLIPRAGWLAGICDPPL 120 | | |
| DB | 61 PVLGLVCPPLGASADHMRGRGRRPRITWALSGLISLFLIPRAGWLAGICDPPL 120 | | |
| QY | 121 ELATLILGVLDFGCGVCFPLEALLSLDFPDPHGQAVSVFMSTIGCGLYLPA 180 | | |

Db 121 ELALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGGCGLYLLPA 180
QY 181 IDWTSALAPYLQTOECLFGLTLFLTCVAATLLVAEEAALGTEPAEGLSAPLSPH 240
Db 181 IDWTSALAPYLQTOECLFGLTLFLTCVAATLLVAEEAALGTEPAEGLSAPLSPH 240
QY 241 CCPCRARLAFRNIGALLPRLHQLCCRMPTLRRLFEVLELCSMMALMTFTLEFYDFVGEGL 300
Db 241 CCPCRARLAFRNIGALLPRLHQLCCRMPTLRRLFEVLELCSMMALMTFTLEFYDFVGEGL 300
QY 301 YQGVPRAPGTEARRHDEGVMSLGLFLQCAISLVFSYVMDRLVQRFGRAYVLA SVA 360
Db 301 YQGVPRAPGTEARRHDEGVMSLGLFLQCAISLVFSYVMDRLVQRFGRAYVLA SVA 360
QY 361 AFPVAAAGATCISHAVVYVTAASALTGFTFSALQILPTTLASLYHREKOVFLPKYRGDTGG 420
Db 361 AFPVAAAGATCISHAVVYVTAASALTGFTFSALQILPTTLASLYHREKOVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPAPFPNGHVAGSGGLPPPALCGASACDVSVRVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPAPFPNGHVAGSGGLPPPALCGASACDVSVRVVGEPTEA 480
QY 481 RVVPGRGICLDLALIDSFAFLISQVAPSLFMGSIYQLSGSVTAYVMSAAGLGLVAITYEATQ 540
Db 481 RVVPGRGICLDLALIDSFAFLISQVAPSLFMGSIYQLSGSVTAYVMSAAGLGLVAITYEATQ 540
QY 541 VFEDKSDIAKYS 553
Db 541 VFEDKSDIAKYS 553

RESULT 2
Q95K15 PRELIMINARY: PRT: 553 AA.
AC Q95K15.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEICAL 59.4 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TEMPORAL LOBE RIGHT;
RA Suzuki Y., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB060851; BAB46871.1; -
KW Hypothetical protein.
SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91FB3BF1E CRC64;

Query Match 97.8%; Score 2798; DB 6; Length 553;
Best Local Similarity 98.2%; Pred. No. 4e-208;
Matches 542; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 MWQRLVWSRLRRRAQLLVNLTFFGLEVCLAAGITVVPPLLEVEEKEFMVNLGIG 60
Db 1 MWQRLVWSRLRRRAQLLVNLTFFGLEVCLAAGITVVPPLLEVEEKEFMVNLGIG 60
QY 61 PVLGLVCPVPLGASDHWKRGYGRRRPFIWALSGLILSLFLIPRAGWLAGLCPDPRPL 120
Db 61 PVLGLVCPVPLGASDHWKRGYGRRRPFIWALSGLILSLFLIPRAGWLAGLCPDPRPL 120
QY 121 ELALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGGCGLYLLPA 180
Db 121 ELALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGGCGLYLLPA 180

QY 181 IDWTSALAPYLQTOECLFGLTLFLTCVAATLLVAEEAALGTEPAEGLSAPLSPH 240
Db 181 IDWTSALAPYLQTOECLFGLTLFLTCVAATLLVAEEAALGTEPAEGLSAPLSPH 240
QY 241 CCPCRARLAFRNIGALLPRLHQLCCRMPTLRRLFEVLELCSMMALMTFTLEFYDFVGEGL 300
Db 241 CCPCRARLAFRNIGALLPRLHQLCCRMPTLRRLFEVLELCSMMALMTFTLEFYDFVGEGL 300
QY 301 YQGVPRAPGTEARRHDEGVMSLGLFLQCAISLVFSYVMDRLVQRFGRAYVLA SVA 360
Db 301 YQGVPRAPGTEARRHDEGVMSLGLFLQCAISLVFSYVMDRLVQRFGRAYVLA SVA 360
QY 361 AFPVAAAGATCISHAVVYVTAASALTGFTFSALQILPTTLASLYHREKOVFLPKYRGDTGG 420
Db 361 AFPVAAAGATCISHAVVYVTAASALTGFTFSALQILPTTLASLYHREKOVFLPKYRGDTGG 420
QY 421 ASSEDSLMTSFLPGPKPAPFPNGHVAGSGGLPPPALCGASACDVSVRVVGEPTEA 480
Db 421 ASSEDSLMTSFLPGPKPAPFPNGHVAGSGGLPPPALCGASACDVSVRVVGEPTEA 480
QY 481 RVVPGRGICLDLALIDSFAFLISQVAPSLFMGSIYQLSGSVTAYVMSAAGLGLVAITYEATQ 540
Db 481 RVVPGRGICLDLALIDSFAFLISQVAPSLFMGSIYQLSGSVTAYVMSAAGLGLVAITYEATQ 540
QY 541 VFEDKSDIAKYS 552
Db 541 VFEDKSDIAKYS 552

RESULT 3
Q95K15 PRELIMINARY: PRT: 501 AA.
AC Q95K15.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEICAL 53.4 KDA PROTEIN.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheidae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MEDULLA OBLONGATA;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RT "Isolation of full-length cDNA clones from macaque brain cDNA
libraries."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AB062977; BAB60745.1; -
KW Hypothetical protein.
SQ SEQUENCE 501 AA; 53447 MW; 8C54ABBD04E0470 CRC64;

Query Match 88.5%; Score 2531; DB 6; Length 501;
Best Local Similarity 97.8%; Pred. No. 1.5e-187;
Matches 489; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 53 MTNVLGTPVGLGVCPVPLGASDHWKRGYGRRRPFIWALSGLILSLFLIPRAGWLAGL 112
Db 53 MTNVLGTPVGLGVCPVPLGASDHWKRGYGRRRPFIWALSGLILSLFLIPRAGWLAGL 112
QY 113 LCPDPRPLEALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGG 172
Db 113 LCPDPRPLEALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGG 172
QY 61 LCPDPRPLEALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGG 120
Db 61 LCPDPRPLEALLIGVLLDFCGGVCFTPLEALLSDLEFRDPDHCROAVSYVAFMISLGG 120
QY 173 CLGYLLPAIDWTSALAPYLQTOECLFGLTLFLTCVAATLLVAEEAALGTEPAEGL 232
Db 173 CLGYLLPAIDWTSALAPYLQTOECLFGLTLFLTCVAATLLVAEEAALGTEPAEGL 232
QY 233 SAPLSPPCCPCRARLAFRNIGALLPRLHQLCCRMPTLRRLFEVLELCSMMALMTFTLEFY 292
Db 233 SAPLSPPCCPCRARLAFRNIGALLPRLHQLCCRMPTLRRLFEVLELCSMMALMTFTLEFY 292

Jun 27 09:46:44 2002

Db 181 SAPSLPSHCCPCWARIAPFNLCALLPRLHQLCCMRPRLRLRLVAELCSWMLMTPTLY 240
 QY 293 TPEVGGGLTQVGPRAEPGRRHRHDEGVMSGLGLFLOCAISLVSFMDRLVORFGR 352
 Db 241 TPEVGGGLTQVGPRAELGTEARRHDEGRMSGLGLFLOCAISLVSFMDRLVORFGR 300
 QY 353 AVLASVAAPVAAGATCLSHSAVAVTASALTGFTSALQILPYTLASLYHREKOVFLP 412
 Db 301 AVLASVAAPVAAGATCLSHSAVAVTASALTGFTSALQILPYTLASLYHREKOVFLP 360
 QY 413 KYRGDTGCGASSEDLSMTSLPGPKPGAPFPNGHVGAGSGLLPPPALCGASACDVSVRY 472
 Db 361 KYRGDGGTSSSEDSMTSLSPGPKPGAPFPNGHVGAGSGLLPPPALCGASACDVSVRY 420
 QY 473 VVEEPPEARVPPGIGICLDLAIIDSAPFLISQVAPSLFMGSIYQLSQSVATYVYSAAGLGL 532
 Db 421 VVEEPPEARVPPGIGICLDLAIIDSAPFLISQVAPSLFMGSIYQLSQSVATYVYSAAGLGL 480
 QY 533 VAIFYAFQVVEFKSDIAKYS 552
 Db 481 VAIFYAFQVVEFKSDIAKYS 500

RESULT 4
 QYSV1 PRELIMINARY; PRT; 599 AA.
 ID 09VSV1
 AC 09VSV1
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
 DE CG4484 PROTEIN.
 GN CG4484.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 NC NCBI_TaxID=7227;
 RX MEDLINE=20196006; PubMed=10731132;
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton R.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
 Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jajani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 Palazzolo M., Peterson G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Sidenkimas I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003552; AAF50310.1;
 DR FlyBase: FBgn0035968; CG4484.
 SQ SEQUENCE 599 AA; 66057 MW; C5381D334CFBF2EB CRC64;

Query Match 18.2%; Score 519.5; DB 5; Length 599;
 Best Local Similarity 23.7%; Pred. No. 4,7e-32;
 Matches 150; Conservative 100; Mismatches 201; Indels 183; Gaps 14;

QY 12 RHRKAQLLVNLTGLGVLCAAGTTPPLLEVGVEEKKMTWVIGIPILGCVPL 71
 Db 46 RRTREMFELSAIAMAIFAAETSFVSPILQIGVHKHMSMTWGLSPILGFPMSPL 105
 QY 72 GSASDHMGRCYGRRRPPTWALSGLISLFLIP-----R 105
 Db 106 GSISDRCKLRMRPPIISILSFGIMCGILVPGKDLGLGAGYTAASALNFTSS 165
 QY 106 AGWLAGLGC-----PDRPELEA--LLILGVGLDPCGVCFPELEALSDLPDPHC 157
 Db 166 GGSVALVSGEATGPGSADYKFAVILITLIGVILGDPADTQTPARTYLLDMC-VPEEQ 224
 QY 158 RQASVYAAWMSLGGCLGLALPDMDSFALAPYGLDCECFGLTLITFCVAAAT-- 215
 Db 225 PRMTWFAFPAGGGTITGALGVMDETHHISFMGNGIPIYFLVITFIANCYLIITYT 284
 QY 216 -----LVAEEAALGP-----TEPABG 231
 Db 285 FREIPLPLEQBELNLPSEQAIRKELKKNNITYIETFLELOMASDDPKREALQG 344
 QY 232 LSAPSLSP-----HCCPRARAFENIGALLPRLHQLCCMRPRLRL 275
 Db 345 SYONGYSPAVEKQKSQDLETOSDYDAPVSLKYLKISITF-----HPSMRMLA 393
 QY 276 VAEICSMMLMTPTLFTYDPEGLYQVGPRAEPGRRHRHDEGVMSGLGLFLOCAIS 335
 Db 394 LTNLFCMGNVTCYCLFTDVGCAVPHGDTAPNSEALNRYBAGVRCGMAIYAFSC 453
 QY 336 LVSLSVMDRLVORFGRRAVYASVAAPV---AAGATCLSHSAVAVTASALTGFTSAL 392
 Db 454 SIYSLVATKMKMKGKRAVYISGMIIYIGIMLVGMPRTKMGVAVSTSA---GLYGTI 510
 QY 393 QILPYTLASLYHREKOVFLPKYRGDGGASSEDLSMTSLPGPKPGAPFPNGHVGAGSG 452
 Db 511 FYVPEFLVARYH-AKNCFSIK-----NG----- 532
 QY 453 LPPPALCGASACDVSVRYVGEPTPEARVP--GRGICLDLAIIDSAPFLISQVAPSLF 509
 Db 533 -----ELVPLKQARGIGDTVAIISSWFIQILIVSLS 564
 QY 510 MGSIVQLSQSVATYVYSAAGLVAIFYAQVYF 543
 Db 565 VGPLVSMDDTCAVLVASTFLSLAALAAAFVLY 598

RESULT 5
 ID 090274 PRELIMINARY; PRT; 576 AA.
 AC 090274
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE MEMBRANE-ASSOCIATED TRANSPORTER PROTEIN B.
 OS Oryzias latipes (Medaka fish).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopteleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percormorpha; Atherinomorpha;
 OC Belontiiformes; Atherinichthyidae; Oryziatidae; Oryzias.

| | | | | |
|-----------------------|------------------|-------------------|------------|------------|
| Query Match | 15.98 | Score 456 | DB 13 | Length 576 |
| Best Local Similarity | 24.08 | Pred. NO. 3.6e-27 | | |
| Matches 144 | Conservative .95 | Mismatches 198 | Indels 164 | Gaps 17 |

| RESULT | 6 | |
|--------|-----------------------|--|
| 065803 | | |
| ID | 065803 | PRELIMINARY; |
| AC | 065803; | PRT; 515 AA. |
| DT | 01-AUG-1998 | (TIREMBLrel. 07, Created) |
| DT | 01-AUG-1998 | (TIREMBLrel. 07, Last sequence update) |
| DT | 01-DEC-2001 | (TIREMBLrel. 19, Last annotation update) |
| DT | SUCROSE/H+ S1MPORTEr. | |

OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids II; Apiales; Apiaceae; Daucus.
OX NCBI_TaxId=4039;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CV, NAMEWISE; TISSUE-ROOT;
RX MEDLINE=99063785; PubMed=9847123;
RA Shakra R., Sturm A.;
RT "Characterization of source- and sink-specific sucrose/H+ symporters
from carrot."; 118:1473-1480(1998).
RL Plant Physiol. 118:1473-1480(1998).
CC -|- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -|- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: Y16768; CAA73659.1;-
DR InterPro: IPR003662; sub.transporter.
DR Pfam: PF00083; sugar_trr_1.
KW Transmembrane.
SQ SEQUENCE 515 AA; 54424 MW; 897B90657C9E243C CRC64;

| | Query Match | 12.1%; | Score 347.5; | DB 10; | Length 515; |
|----|--|--------------|--------------------|-----------------|---------------------|
| | Best Local Similarity | 25.9%; | Pred. No. 7.7e-19; | | |
| | Matches 131; | Conservative | 80; | Mismatches 215; | Indels 79; Gaps 17; |
| OY | 17 QLLNVNLTFFGLEVCVLAAGITVPPLLEVEGEKFMVTWLGIGPVLCVCPILGSASD | 76 | : | : :: : | :: : |
| Dd | 34 KLVVAIAIAGVGFGMALQLSLTTPYQGLGIPHMAVAIMICGISGMIVGPVIYGYS | 93 | : | : :: : | :: : |
| OY | 77 HNRGRGRRRPTMALSGILLSLFLIRPAGLAGI-----LCDDPRPLEALLII | 129 | : | : :: : | :: : |
| Dd | 94 HCQSSEGRRRRPIASCAGCVAILSVLI--GFADISYKAGDDMSKTLPRAVFVVIGF | 150 | : | : :: : | :: : |
| OY | 130 GLDFCGGYCFPLBALSDLFR-DPDHCROYAVYAEMISLGCLG-----LL | 178 | : | : :: : | :: : |
| Dd | 151 WLIDVANNMLQSPCRALLADCSGPTRRKRSMATYSEPMANGNLTAAGSYNNLYLF | 210 | : | : :: : | :: : |
| OY | 179 PAIDMDSALAPYLCTOEBCFLGLTLFLFCVAATLLVAEEAALGPEPAEGISAPSLS | 238 | : | : :: : | :: : |
| Dd | 211 PPS-KTHACDIYCANKLSCEFIISIALIIITVVALSVRENS-GPPDDADAEEP--- | 263 | : | : :: : | :: : |
| OY | 239 PHCCPFRAFLARNIGALLPRFHOCSCMPTRLRLFAVELCSMMALMTFLFTDFGE | 298 | : | : :: : | :: : |
| Dd | 264 ----FSSGRIIP-FGELLGALKDL----PRMLLLIIVTCINMTAMFPFLTFDTMNGR | 313 | : | : :: : | :: : |
| OY | 299 GLYGVPRAPEGTEARRHYDEGVRMGSLGPLQCAISLVSLVMDRLVORETRAVYLAS | 358 | : | : :: : | :: : |
| Dd | 314 ELYGT--AGQG---KLVDGVRAGALGILLINSVGLTSLAVEYIVRGVGYKI-LMG | 366 | : | : :: : | :: : |
| OY | 359 VAAPFYAARGACL-----SHSV-----AYTTASA-----ALTGETTSALQIL | 395 | : | : :: : | :: : |
| Dd | 367 FYNFLTALIGLVTVVSVKVAOHQEHSHANSGOLLPPSACVAKGALSLEFILIPLSITYSI | 426 | : | : :: : | :: : |
| OY | 396 PYTLIASLHYRREKYVFLPKYRQDTGGASSEDLSMTSFLEGKPRG-----APPFNCHGYA | 448 | : | : :: : | :: : |
| Dd | 427 PFALASIYSSGGAQGLSIGLVNALAIYVPMIYSVLAAPPDSLFGGNLPAIFYGAISA | 486 | : | : :: : | :: : |
| OY | 449 GGSGLP-----PPPALCASACDVS 469 | | : | : :: : | :: : |
| Dd | 487 AISGLVALIVLLPKDSKAASKLSLS 511 | | : | : :: : | :: : |

```

RESULT      7
O9FNRR6    PRELIMINARY;          PRT;         515 AA.
AC O9FNRR6;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE SUCROSE/PROTON SYMPORTER.
OS SUT2.
GC Daucus carota (Carrot).
NC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

```

| | | | |
|--------|--|---|--------------|
| OY | 17 | QLLLVNLTFEGEVLCAAGITVVYPPLLELVGVEEFMTMVLIGIPYGLVGVLPGLSGASD | 76 |
| : | : | : | : |
| Dd | 34 | KLVVAIAAAGVOFGMALQDLSTLTPYVDLLGPBHWAAVYMLCGISOMLYQPITYGYSD | 93 |
| OY | 77 | HMGGRGRRRPFVWALSLGIILSLFLPACWLGL-----LCPPRLPEALLILICV | 129 |
| : | : | : | : |
| Dd | 94 | HCSSFRRRRPFIASAGCVAISVII---GRAADISKAGDDMSKTLPRAVTVFVIGE | 150 |
| OY | 130 | GLDFCGOVCFFPLEALLSDLFR-DEPHCRQAYSVYAEMISLGCLGY-----LL | 178 |
| : | : | : | : |
| Dd | 151 | WIIDVANNNMLOGPCRALIDLDCSGDTRRMRSNAFYSEFMGAANGTILGTAAAGSYNNLYRLF | 210 |
| OY | 179 | PAIDMDTSALAPLIGTOBECLFGLLILFELTCVAATILVAEEBAALGPTEPAIGLSAPSLS | 238 |
| : | : | : | : |
| Dd | 211 | PFS--KTHADDLCANLKSCFTTISTALLIITYVALSVRENS-GPPDDADAEEP---- | 263 |
| OY | 239 | PHCCPRCARLAFTNGLALPRLHOLCCMRPTLRRLFVAELCSWMALMFIFYDYDVE | 298 |
| : | : | : | : |
| Dd | 264 | ----PSSGRIPV-FGLLGLAKDL---DRPMILLIIVTCLMWIMAFPIILDIDMMGR | 313 |
| OY | 299 | GLYGQVPRAEPGETARRRHEDGVAMSGSLGFLOCAISLVSLEYMDRLVORFGTAVYLAS | 358 |
| : | : | : | : |
| Dd | 314 | EYGGT--AGKG---KLVDGVRAGAULGLLNSVYGLITSINAEVLTNRGVGVKI-LWG | 366 |
| OY | 359 | VAAFVYVANGARCL-----SHSV-----AVVTASA---ALTGFTFSALQILI | 395 |
| : | : | : | : |
| Dd | 367 | FYNFLAIAGLWNTVYVSKVAQHOREHSANGOLLPPSAGVAGALSIFSLIGLSTYST | 426 |
| OY | 396 | PYTLASLVYHREKQVFLLPYRGDTGASSEDLSIMSPFGPPG-----APPNGHVGA | 448 |
| : | : | : | : |
| Dd | 427 | PALASITSYSGSAGOCISLGLNALTAIVPOMITVASLAGPDSLFGCGNLPAFVVGALISA | 486 |
| OY | 449 | GGSGLP-----PPALCGASACDVS 469 | |
| : | : | : | : |
| Dd | 487 | AISGVLAIVLLPKPSDAASKSLSS 511 | |
| RESULT | 8 | | |
| ID | O9SOK6 | PRELIMINARY; | PRT; 501 AA. |
| AC | O9SOK6; | | |
| Dt | 01-MAY-2000 (TREMBREL. 13, Created) | | |
| Dt | 01-MAY-2000 (TREMBREL. 13, last sequence update) | | |
| Dt | 01-DEC-2001 (TREMBREL. 19, last annotation update) | | |
| DE | POTATIVE SUCROSE TRANSPORTER. | | |
| GN | VYSUC11. | | |
| OS | Vitis vitifera (Grape). | | |
| OC | Eukaryotes: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons: core eudicots; Vitaceae; | | |
| OC | Vitis. | | |
| OX | NCBL_taxid=29760; | | |
| RN | [1] | | |

| | | | | |
|---------------------------|--------|--------------------|-------------|-------------|
| Query Match | 12.1%; | Score 346.5; | DB 10; | Length 501; |
| Best Local Similarity | 27.0%; | Pred. No. 8.9e-19; | | |
| Matches 153; Conservative | 83; | Mismatches 197; | Indels 133; | Gaps 22 |

[illegible]

| Q9SLN7 | PRELIMINARY; | PRT; | 515 AA. |
|--------|---|------|---------|
| OC | Q9SLN7 | | |
| AC | Q9SLN7; | | |
| DT | 01-MAY-2000 (TREMBLrel. 13, Created) | | |
| DT | 01-MAY-2000 (TREMBLrel. 13, Last sequence update) | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) | | |
| DE | SUCROSE TRANSPORTER PROTEIN. | | |
| GN | CSUT. | | |
| OS | Daucus carota (Carrot). | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| OC | Asteridae; eusterids II; Apiales; Apiaceae; Daucus. | | |

RC STRAIN=SHIRAZ; TISSUE=FRUIT;
RA Davies C., Wolf T., Robinson S.P.;
RT "three putative sucrose transporters are differentially expressed in
RT RT grapevine tissues.";
RL Plant Sci. 147:93-100(1999).
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY)
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL: AF021809; AAF08330.1; -;
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugar_tr. 1.
KW Transmembrane.
SQ SEQUENCE 612 AA; 65711 MW; 8FC3FED6A8439078 CRC64;

| Query Match | 12.18; | Score 346; | DB 10; | Length 612; |
|-----------------------|------------------|--------------------|-------------|-------------|
| Best Local Similarity | 25.48; | Pred. No. 1.2e-18; | | |
| Matches 125; | Conservative 75; | Mismatches 178; | Indels 114; | Gaps 177; |

[illegible]

| ID | Q9S7Z5 | PRELIMINARY: | PRT: | 512 AA. |
|----|---|--------------|------|---------|
| AC | Q9S7Z5; | | | |
| DT | 01-MAY-2000 (TREMBREL, 13, Created) | | | |
| DT | 01-MAY-2000 (TREMBREL, 13, Last sequence update) | | | |
| DT | 01-OCT-2001 (TREMBREL, 18, Last annotation update) | | | |
| DE | SUCROSE TRANSPORTER SUT2B. | | | |
| GN | SUT2B OR SUT2A. | | | |
| OS | Apium graveolens (Celery). | | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta | | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | | |
| OC | Asasteridae; euasterids II; Apiales; Apiaceae; Apium. | | | |
| OX | NCBI_taxid=4045; | | | |
| RN | [1] | | | |

```

RP SEQUENCE FROM N.A.
RC TISSUE=MATURE LEAF;
RA Noiraud N., Lemoine R.;

```


OC euroids ¹, Fabales; Fabaceae; Papilionoideae; Viciaeae; Vicia.
 OX NCBI_TaxID=3906;
 RN [1]

UC
VLIS:
OX NCBI_TaxID=29760;
RN [1]
RP SEQUENCE FROM N.A.

| | | |
|--------|--|---------------------------|
| RESULT | 13 | |
| Q9SP63 | | |
| ID | Q9SP63 | PRELIMINARY; PRT; 501 AA. |
| AC | Q9SP63; | |
| DT | 01-MAY-2000 (TREMblrel. 13, Created) | |
| DT | 01-MAY-2000 (TREMblrel. 13, last sequence update) | |
| DT | 01-DEC-2001 (TREMblrel. 19, last annotation update) | |
| DE | SUCROSE TRANSPORTER. | |
| OS | Vitis vinifera (Grape). | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Vitaceae; | |
| OC | Vitis. | |
| OX | NCBI_TaxID=29760; | |
| RN | [1] | |
| RP | SEQUENCE FROM N.A. | |
| RA | STRAIN=CV. Ugni Blanc; TISSUE=GRAPE BERRY; | |
| RA | Accession A. Tressaly N., Picaud S., Delrot S., Romieu C.; | |

| | | | |
|--------|---|---|--------------|
| RESULT | 14 | | |
| 09ZTB9 | | | |
| ID | 09ZTB9 | PRELIMINARY; | PRT; 512 AA. |
| AC | 09ZTB9 | | |
| DT | 01-MAY-1999 | (TREMBlrel. 10, Created) | |
| DT | 01-MAY-1999 | (TREMBlrel. 10, Last sequence update) | |
| DT | 01-DEC-2001 | (TREMBlrel. 19, Last annotation update) | |
| DE | SUCROSE TRANSPORTER. | | |
| GN | SUT1. | | |
| OS | Apium graveolens (Celery). | | |
| OC | Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| OC | Asteridae; euasterids II; Apiales; Apiaceae; Apium. | | |
| NCBI | taxid=4045; | | |

[illegible]

| | | | |
|--------|--|--------------|--------------|
| RESULT | 15 | | |
| 043653 | | | |
| ID | Q43653 | PRELIMINARY; | PRT; 516 AA. |
| AC | Q43653; | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Created) | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Last sequence update) | | |
| DT | 01-DEC-2001 (TREMBLrel. 19, Last annotation update) | | |
| DE | SUCROSE TRANSPORT PROTEIN. | | |
| GN | SUT1. | | |
| OS | Solanum tuberosum (potato). | | |
| OC | Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; | | |
| OC | Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; | | |
| OC | Asteridae; euasterids I; Solanales; Solanaceae; Solanum. | | |
| OX | NCBI_TaxID=4113; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |

RC STRAIN-CV DESIRE: TISSUE-LEAF;
 MEDLINE-94146554; PubMed-8312741;
 RA Riesmeier J., Willmitzer L., Frommer W.B.;
 RT "Potato sucrose transporter expression in minor veins indicates a role
 in phloem loading."
 RL Plant Cell 5:1591-1598(1993).
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 DR EMBL: X69165; CAA48915.1; -
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transmembrane.
 SQ SEQUENCE 516 AA; 54831 MW; 4FD06C095E49A377 CRC64;

Query Match 11.5%; Score 330; DB 10; Length 516;
 Best Local Similarity 24.7%; Pred. No. 1.7e-17;
 Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

QY 4 RLWVSRLLHRRKQALLVLTFFGLVLCIAGITVPPILLLEVGEKEMTWIGIPVL 63
 DB 29 KLM-----KIVASIAAGVGFALQLSLTPYQLGIPHKFASFIWLCGPIS 78
 QY 64 GLVCPVLLGSADHWGRYGRRRPTWALSGLISLPIPRAGWLAGL-----LCPD 116
 DB 79 GMIVQPVYGYSDNCSSRFGRRRPFIAGALVMTAVFLI--GEADIGHASGDTLGKG 135
 QY 117 PRLELALLIGVGLDFGVCYCFPLEALLSDLEFRPD-HCRQAVSYAFMISLGCLG 175
 DB 136 FKPRALAVFVGVWILDVANNMLOGPCRALLDLSGKSGRMRTANAFSFMVAGNLLG 195
 QY 176 YLLPAIDW-----DTSALAPYLCTOECLF-GILLIFLTCVATLLVAE-AALGP 225
 DB 196 YAGSSYHLEKVPFSKTKACDWCANLKSCEFIATIFLLSITITLITVRENELPEKDE 255
 QY 226 TEPAGEISAPSLSPCCPCARLAFNLGALLPRHLQCCMRPTLRLEVAELCSMAL 285
 DB 256 QETDEKLAGAG-----KSKVPF--FGIFGALKEL--PRPWILLIYTCNMIAW 301
 QY 286 MFTLFYTFVGEGLYGVPRAPGTEARRHYDEGVMSGLFLQCAISLVFSLVMDRL 345
 DB 302 FPFPLIYTDMAKEVEGG---QVGD--ARLYDLGVAGAMGILLQSVVLGFMSLGEFL 355
 QY 346 VQREGTRAVYLVAAFPVAAGATCLSHSAVAV-----TASALTGFT----- 388
 DB 356 GKRIQ-GAKRLMGILNFVL--AICLAMTILYTKMAEKSRQHPDAGTLMGPTPGKIGAL 411
 QY 389 --PSALQI-----LYTTIASLVHREK-----QVFLPKYRGDTGASSEDS 426
 DB 412 LLEFALGIPPLAATRSIFPALASTISSNRGSGGLSLGVNLMAIVPQMLVSLVGGPMDL 471
 QY 427 LMTSFLPGPKGAPFPNGHWAGSGS-----LPPPALCGASACDVSVRYVYG 475
 DB 472 FGGGNLPG-----FVVGAVAAASAVLALTLMPSPPA-----DAKPAVAMG 512

Search completed: June 26, 2002, 13:47:28
 Job time: 611 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:41:37 ; Search time 37.41 Seconds
(without alignments)
572.358 Million cell updates/sec

Title: US-09-838-785-2
Perfect score: 2861
Sequence: 1 MVGRLLWVSRLLRRKQALL.....AIYFATQVPEKSLAKTSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 489.5 | 17.1 | 530 | 1 | MARP_MOUSE |
| 2 | 488.5 | 17.1 | 530 | 1 | MARP_HUMAN |
| 3 | 324.5 | 11.3 | 525 | 1 | STP_SPTOL |
| 4 | 142 | 5.0 | 544 | 1 | YD74_SYNY3 |
| 5 | 136 | 4.8 | 541 | 1 | G710_HUMAN |
| 6 | 125.5 | 4.4 | 539 | 1 | TCR1_ECOLI |
| 7 | 121.5 | 4.2 | 503 | 1 | PUR8_STRLP |
| 8 | 119.5 | 4.2 | 491 | 1 | AMPG_ECOLI |
| 9 | 117 | 4.1 | 440 | 1 | HJTE_ECOLI |
| 10 | 117 | 4.1 | 495 | 1 | SWVA_SALTY |
| 11 | 111.5 | 3.9 | 640 | 1 | Y051_MCTU |
| 12 | 111 | 3.9 | 473 | 1 | PHDK_NCCSK |
| 13 | 110.5 | 3.9 | 680 | 1 | CAIA_HUMAN |
| 14 | 110 | 3.8 | 368 | 1 | GALT_HUMAN |
| 15 | 110 | 3.8 | 606 | 1 | NOOC_THERH |
| 16 | 109.5 | 3.8 | 618 | 1 | SPH2_HUMAN |
| 17 | 109 | 3.8 | 477 | 1 | TPUM_RHOCA |
| 18 | 107 | 3.7 | 419 | 1 | CMLA_PSEAE |
| 19 | 105 | 3.7 | 448 | 1 | YJ94_YEAST |
| 20 | 104 | 3.6 | 476 | 1 | MELB_SALTY |
| 21 | 103 | 3.6 | 465 | 1 | FTSM_MYCLE |
| 22 | 102.5 | 3.6 | 461 | 1 | PUCG_RHOCA |
| 23 | 102 | 3.6 | 438 | 1 | SHIA_ECOLI |
| 24 | 101.5 | 3.5 | 428 | 1 | YX10_BACSU |
| 25 | 101 | 3.5 | 437 | 1 | BRAZ_PSEAE |
| 26 | 101 | 3.5 | 471 | 1 | MELB_PSEAE |
| 27 | 101 | 3.5 | 471 | 1 | MELB_KLEPN |
| 28 | 100 | 3.5 | 404 | 1 | YAYM_RHISN |
| 29 | 100 | 3.5 | 473 | 1 | YIHO_SALTY |
| 30 | 99 | 3.5 | 481 | 1 | LMRA_STRLN |
| 31 | 98.5 | 3.4 | 613 | 1 | FNOL_ECOLI |
| 32 | 98.5 | 3.4 | 685 | 1 | FHUB_SALTY |
| 33 | 98 | 3.4 | 379 | 1 | CYB_SORCI |

| | | | | | | |
|----|------|-----|------|---|------------|---------------------|
| 34 | 98 | 3.4 | 496 | 1 | GTR3_CHICK | P28568 gallus gall |
| 35 | 97.5 | 3.4 | 354 | 1 | RNPD_PSEST | Q9evn4 pseudomonas |
| 36 | 97.5 | 3.4 | 385 | 1 | P12R_BOVIN | P79393 bos taurus |
| 37 | 97.5 | 3.4 | 433 | 1 | DCUA_WOLIN | O34245 wolfinella s |
| 38 | 97.5 | 3.4 | 1103 | 1 | CYGD_HUMAN | O02846 homo sapien |
| 39 | 97 | 3.4 | 336 | 1 | CYB_SORGA | O79966 sorax grana |
| 40 | 97 | 3.4 | 336 | 1 | CYB_SORMO | O79452 sorax hayde |
| 41 | 97 | 3.4 | 336 | 1 | CYB_SORNO | O79969 sorax monti |
| 42 | 97 | 3.4 | 336 | 1 | CYB_SORVA | O80019 sorax vagta |
| 43 | 97 | 3.4 | 370 | 1 | GALT_RAT | O88626 ratus norv |
| 44 | 97 | 3.4 | 427 | 1 | RBT_KLEPN | O52717 klebsiella |
| 45 | 96.5 | 3.4 | 400 | 1 | TCR8_PASMU | P51564 pasteurella |

ALIGNMENTS

| RESULT | ID | Sequence | Standard | PRT | 530 AA. |
|--------|--|-----------|----------|---------|---------|
| 1 | MARP_MOUSE | STANDARD; | PRT; | 530 AA. | |
| AC | P58355; | | | | |
| DT | 01-MAR-2002 (Rel. 41, Created) | | | | |
| DT | 01-MAR-2002 (Rel. 41, Last sequence update) | | | | |
| DT | 01-MAR-2002 (Rel. 41, Last annotation update) | | | | |
| DE | Membrane-associated transporter protein (AIM-1 protein) (Melanoma | | | | |
| DE | antigen AIM1) (Underwhite protein). | | | | |
| GN | MARP OR AIM1 OR UW. | | | | |
| OS | Mus musculus (Mouse). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| OX | NCBI_TaxID=10090; | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RP | STRAIN-DDY; TISSUE-Eye, Kidney, and Uterus; | | | | |
| RC | MEDLINE=21372467; PubMed=11479596; | | | | |
| RA | Fukamachi S., Shimada A., Shima A.; | | | | |
| RT | "Mutations in the gene encoding B, a novel transporter protein, reduce | | | | |
| RT | melanin content in medaka."; | | | | |
| RL | Nat. Genet. 28:381-385(2001). | | | | |
| RL | [2] | | | | |
| RP | SEQUENCE FROM N.A., AND VARIANTS UW-DBR ASN-153 AND PRO-435. | | | | |
| RX | MEDLINE=21473748; PubMed=11574907; | | | | |
| RA | Newton J.M., Cohen-Barak O., Hagiwara N., Gardner J.M., Davison M.T., | | | | |
| RA | King R.A., Brilliant M.H.; | | | | |
| RT | "Mutations in the human orthologue of the mouse underwhite gene (uw) | | | | |
| RT | underlie a new form of oculocutaneous albinism, OCA4."; | | | | |
| RL | Am. J. Hum. Genet. 69:981-988(2001). | | | | |
| CC | -1- FUNCTION: Melanocyte differentiation antigen. May transport | | | | |
| CC | substances required for melanin biosynthesis (By similarity). | | | | |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By | | | | |
| CC | similarity). | | | | |
| CC | -1- TISSUE SPECIFICITY: Melanocytes, eyes, kidney and uterus. | | | | |
| CC | -1- DISEASE: Defects in MARP are the cause of the uw-dbr phenotype | | | | |
| CC | that results in loss of nearly all pigmentation in the homozygous | | | | |
| CC | state. | | | | |
| CC | -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. | | | | |
| CC | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | | |
| CC | or send an email to license@isb-sib.ch). | | | | |
| DR | EMBL: AF360357; AAK81713.1; | | | | |
| KW | Melanin biosynthesis; Transmembrane; Antigen; Glycoprotein; Vision; | | | | |
| KW | Disease mutation; Albinism. | | | | |
| FT | DOMAIN 1 45 | | | | |
| FT | DOMAIN 46 66 | | | | |
| FT | DOMAIN 67 68 | | | | |
| FT | TRANSMEM 69 89 | | | | |
| FT | FT 90 105 | | | | |
| FT | DOMAIN | | | | |
| FT | CYTOPLASMIC (POTENTIAL). | | | | |
| FT | EXTRACELLULAR (POTENTIAL). | | | | |
| FT | 2 (POTENTIAL). | | | | |
| FT | CYTOPLASMIC (POTENTIAL). | | | | |

```

FT TRANSMEM 106 126 3 (POTENTIAL).
FT DOMAIN 127 138 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 184 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 237 6 (POTENTIAL).
FT DOMAIN 238 318 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 319 339 7 (POTENTIAL).
FT DOMAIN 340 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 8 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 419 9 (POTENTIAL).
FT DOMAIN 420 425 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 426 446 10 (POTENTIAL).
FT DOMAIN 447 477 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 478 498 11 (POTENTIAL).
FT DOMAIN 499 504 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 505 525 12 (POTENTIAL).
FT DOMAIN 526 530 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 356 356 N-LINKED (GLCNAc. . .) (POTENTIAL).
FT VARIANT 153 153 D -> N (IN UM-DBR).
FT VARIANT 435 435 S -> P (IN UM-DBR).
SQ SEQUENCE 530 AA: 57961 MW: 145260.791609FC CRC64:

```

```

Query Match 17.1%; Score 489.5; DB 1; Length 530;
Best Local Similarity 26.1%; Pred. No. 1.3e-27;
Matches 146; Conservative 85; Mismatches 228; Indels 101; Gaps 12;

```

```

OY 17 QLLVNLTFGLVCLAGTIVPPLLEVEEKEFMVNLGIVGLVPLGLSASD 76
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 34 RLVMHSMAMFGRFVYAEAAVTVFVLLSVGLPKSLYSAMWLSLPLGLQPVVGSASD 93
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 77 HRRGVRGRRPRPMLSLGILLPLPRAGMLAGLCPDRP---LELALLIGVGLD 133
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
DB 94 HCRAMKGRRRPILTLAIMLGMALYLGDAVSAVLANPQKILWALSTTMGVGLD 153
   | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 134 FCGQVCFPLLEALSLDFRDPHCRQAVSVAFMISLGLCYLLPAIDMTSALAPYL 193
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 154 FSADFIDGRKAYLFDVCSHQK-EKGLHYHALTFGFGALGYILGAIIDWHLDRLLG 212
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 194 TOECLFGLTLIFLTCVAATLVAEAL--GPTBP-----AEGLSAPSLSPKCCP 244
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 213 TFEQVWFFFSALVLLCFITHLCSLPEAPLRDAATDPQSDPQSSLSASGMHEX---- 268
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 245 RARLAFRNLA-----LRLQLCCMRPLRLRLFVAFLGSM 282
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 269 GSIEEVKNGADTEQPOVDMKKKPSGSGQRTMSKSLRLALVNMPSHRLCVSHLIG 328
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 283 MALMTFTLEFVDEVEGLYQVPRAPRGTEARRHYDEGVRMGSLGLFLOCAISLVFSLV 342
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 329 TAFLSMFLFTDFMGQIVYHDPYGANHSTERFLIERGEVGCWGLCINSVFSVSYRQ 388
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 343 DLYVDFRGRVYLVSAVPAAGATCLSHSAVAVTASALATGFFSLQLDLPITLASL 402
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 389 KAMVSYTIGKGLYFMGYLLFFGLGTGTGLFPNNVSTLYVCSMFGVASSFLTYVFNLLAE 448
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 403 YAREKOVFLPKYRDPGTGASSEDLSMTSFLPGPKGAPRPNHVGAGSGGLPPRPALG 462
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 449 YIREEE---KEKGEA-----PEGPNQGR---GKV-----DCA 477
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 463 ASACDVSVRVVGEPTAEARVPGRGICLDLALIDSALFLSQVAPSLFMGSIYVOLSQSV 522
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 478 ALTCMVQL-----AQILVGGG-----LQFLVNMAGSYV 506
   : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 523 YWVSAAGLGLVAIYRQYV 542
   : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 507 VVITASAVSLIGCCFVALFV 526
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 2
MATP_HUMAN

```

ID MATP_HUMAN STANDARD: PRT: 530 AA.
AC Q9UWX9; Q9BTM3;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Membrane-associated transporter protein (ATM-1 protein) (Melanoma
DE antigen AIM1).
GN MATP OR AIM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORM AIM-1A).
RP TISSUE=Melanoma;
RX MEDLINE=21115844; PubMed=11221837;
RA Harada M., Li Y.F., El-Gamil M., Rosenberg S.A., Robbins P.F.;
RT "Use of an in vitro immunoselected tumor line to identify shared
RT melanoma antigens recognized by HLA-A*0201-restricted T cells.";
RL Cancer Res. 61:1089-1094(2001).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM AIM-1B).
RP TISSUE=Skin;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN ALTERNATIVE SPLICING.
RA Ferro S.;
RL Unpublished observations (NOV-2001).
RN [4]
RN DISEASE, AND VARIANT LEU-374.
RX MEDLINE=21473748; PubMed=11574907;
RA Newton J.M., Cohen-Barak O., Hagihara N., Gardner J.M., Davison M.T.,
RA King R.A., Brilliant M.H.;
RT "Mutations in the human orthologue of the mouse underwhite gene (uv)
RT underlie a new form of oculocutaneous albinism, OCA4.";
RL Am. J. Hum. Genet. 69:981-988(2001).
CC -1- FUNCTION: Melanocyte differentiation antigen. May transport
CC substances required for melanin biosynthesis (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein; melanosome (By
CC similarity).
CC -1- ALTERNATIVE PRODUCTS: At least 3 isoforms: AIM-1a (shown here),
CC AIM-1b and AIM-1c; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: Expressed in most melanoma cell lines and
CC melanocytes.
CC -1- DISEASE: Defects in MATP are the cause of oculocutaneous albinism
CC type 4 (OCA4). OCA4 is characterized by hypopigmentation of skin,
CC hair and eyes. It leads to reduced visual acuity.
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to a
CC frameshift in position 188.
CC -1- CAUTION: The described alternatively spliced isoforms are inferred
CC using information from ests.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL, AF172849; AAD51812.1;
DR EMBL, BC003597; AA03597.1; ALT_FRAME.
DR MIM, 606202;
KM Melanin biosynthesis; Transmembrane; Antigen; Glycoprotein; Vision;
KW Polymorphism; Albinism; Alternative splicing.
FT DOMAIN 1 46 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 47 67 1 (POTENTIAL).
FT DOMAIN 68 68 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 69 89 2 (POTENTIAL).
FT DOMAIN 90 110 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 111 131 3 (POTENTIAL).

```

```

FT DOMAIN 132 138 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 184 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 216 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 217 237 6 (POTENTIAL).
FT DOMAIN 238 318 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 319 339 7 (POTENTIAL).
FT DOMAIN 340 366 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 367 387 8 (POTENTIAL).
FT DOMAIN 388 398 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 399 419 9 (POTENTIAL).
FT DOMAIN 420 425 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 426 446 10 (POTENTIAL).
FT DOMAIN 447 477 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 478 498 11 (POTENTIAL).
FT DOMAIN 499 504 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 505 525 12 (POTENTIAL).
FT CARBOHYD 526 530 CYTOPLASMIC (POTENTIAL).
FT VARSPLIC 535 356 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPLIC 129 187 MISSING (IN ISOFORM AIM-1C).
FT VARSPLIC 188 295 MISSING (IN ISOFORM AIM-1B).
FT VARSPLIC 386 406 YFOKLVSYIGLKGIFGVL -> CKSFSLRMSSKSFWS
FT VARSPLIC 407 530 /FTID-VAR_012162.
FT VARIANT 374 374 F -> L.
FT SEQUENCE 530 AA: 58301 MW: F14ADACAA8FF31B CRC64:

```

```

Query Match 17.1%; Score 488.5; DB 1; Length 530;
Best Local Similarity 26.4%; Pred. No. 1.5e-27;
Matches 150; Conservative 78; Mismatches 233; Indels 107; Gaps 12;

```

```

OY 8 SLLRHRRAQLLLVLLFPGLEVCIAAGITVYRLLLEVEVEKMTWVIGIPVLGVC 67
DB 33 SLLIHNSAM-----FGREFCYAENAAYTVPVLLSYGLSPSSLSYTWFLSPILGFL 84
OY 68 VPLGSASDHMGRRGRPRFIALSLGILLFLPRAGWAGLCPDR---PLELAL 124
DB 85 QVVGASADHCSRWGRRRPYLLGLVGMALYLGATVVAALLIAPRRKLWALSV 144
OY 125 LILGVLDFGCGVCFPLEALLSDLRDPDRCHQAOAYSVAFMISGLCYLPAIDMD 184
DB 145 TWIGVLEDFEADFTDGRPKATLFDVCSHOK-ENGLHNAHLPFGFGALGVLGALDIMA 203
OY 185 TSALAPVIGTOECLFGLITLFLTCVATLLVAEALGPTEPAEGLSAPLSPHCCPC 244
DB 204 HLELGRILGTEFOVFFESALVLTLCFTVHLCSISEAPL--TEVAKGI-PPQOTPODPL 260
OY 245 RA-----RLAFRNL-----GALLPRLHOLCCMPRTLRL 274
DB 261 SSDGMYEGSTIEKVKNGVNPPELAAQAKNNKHAPOTRRAMTLKSLRLALVMPHYRL 320
OY 275 FVAELCSMMALMTFTLFTYDEVEGLYGVPRABEPTEARRHVDEGVMSGLFLQCAI 334
DB 321 CISHLIGMTAFLSNMLFTTDFMGQIVYRGDPSAHNSTEFLYKGVGCGMGPCINSVF 380
OY 335 SLVSLVMDRLVORFGTAIVLYAABPVVAGATCLSHSAVAVTASALGFTFSALOI 394
DB 381 SLVSYFQKVLVSYIGLKGIFGVLGTLGFLFPNYSVLVLSLFCVMSSTLYT 440
OY 395 LPYTLASLYHREKOVFLKRYGDDTGCGASSEDSLMTSLPDRPGAPFRNGHAGAGSGIL 454
DB 441 VPENLITEYHREE---KERQA-----PGGP----- 465
OY 455 PPPPALCASACDVSRYVVGEPTEARVYVPGRICLDLAIIIDSAFLISQVAPSLFMGSIV 514
DB 466 -----DNSVR-----GKG-MDCATLTCMQVLAQIIVGGGLGLV 498
OY 515 OLSGSVYAVMVSAGGLVAIYFAFOV 542
DB 499 NAGTGVVVVITASAVALLGCCFVALFV 526

```

```

RESULT 3
STP_SPTOL
ID STP_SPTOL STANDARD: PRT: 525 AA.
AC Q03411;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Sucrose transport protein (Sucrose permease) (Sucrose-proton
DE symporter).
OS Spinacia oleracea (spinach).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Caryophyllidae; Caryophyllales; Chenopodiaceae; Spinacia.
OX NCBI_Taxid=3362;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leaf;
RX MEDLINE=93099843; PubMed=1464305;
RA Riesenmeyer J.W., Willmitzer L., Frommer W.B.;
RT "Isolation and characterization of a sucrose carrier cDNA from
RT spinach by functional expression in yeast.";
RL Embo J. 11:4705-4713(1992).
CC -!- FUNCTION: RESPONSIBLE FOR THE TRANSPORT OF SUCROSE INTO THE
CC CELL, WITH THE CONCOMITANT EXPORT OF A PROTON (SYMPORT SYSTEM).
CC CAN ALSO TRANSPORT MALTOSE AT A LESSER RATE.
CC -!- PATHWAY: SUCROSE METABOLIC PATHWAY.
CC -!- SUBCELLULAR LOCATION: INNER MEMBRANE.
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X67125; CAA47604.1; -.
DR PIR: S28052; S28052.
DR InterPro: IPR003662; sub-transporter.
DR Pfam: PF00083; sugar_tr.1.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; FALSE NEG.
DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
KM Transmembrane; transport; Sugar transporter; Symport.
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 38 58 1 (POTENTIAL).
FT TRANSMEM 72 92 2 (POTENTIAL).
FT TRANSMEM 107 127 3 (POTENTIAL).
FT TRANSMEM 145 165 4 (POTENTIAL).
FT TRANSMEM 184 204 5 (POTENTIAL).
FT TRANSMEM 230 250 6 (POTENTIAL).
FT TRANSMEM 295 315 7 (POTENTIAL).
FT TRANSMEM 338 358 8 (POTENTIAL).
FT TRANSMEM 373 393 9 (POTENTIAL).
FT TRANSMEM 422 442 10 (POTENTIAL).
FT TRANSMEM 455 475 11 (POTENTIAL).
FT TRANSMEM 488 508 12 (POTENTIAL).
FT DOMAIN 509 525 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 525 AA: 54992 MW: 018347A4D2C1C6 CRC64;

```

```

Query Match 11.3%; Score 324.5; DB 1; Length 525;
Best Local Similarity 23.2%; Pred. No. 6.4e-16;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

```

```

OY 23 LITFGLVCLAGITV-----VPLLEVEVEKMTWVIGIPVIGLCVPLGSA 74
DB 35 LKKLGIVASVAGVDFGNALQLSLTPYQVLGIFHTMAIYVLCGPISGMTVQPLVGY 94
OY 75 SDHMGRRGRPRFIALSLGILLFLPRAGWLA--GLCPDR-----RPLELALLIL 127
DB 11:||||| | : : : : | | : : : :

```

```

Dh 95 SDRCTSRGRRRPRPLAAGALVAVAVGLI---GPAADIGAASGPTGNVAKPRATIAVEV 151
Qy 128 GVGLLDFCGGYCFFPLPEALSLDLFR-DPDHCROAVSYVAFMISLGCLGYLLPAID---- 182
Dh 152 GFWILLDVANNNTLOGPCCRALADMAAGSOTKTRVYANAFSEFMAIGNGYAAGSYRLYT 211
Qy 183 ----WDTSALAPVYGTQECFLGLLT-LIFLTGYAATLLVAEEAALGPTEPAEGLSAPSL 237
Dh 212 VFPEFTKACACVYCANLKSCEFFISTILLVLTIALSVKEROQTTTIDEIQEEDLKRNN 271
Qy 238 SPHCCPCRCARLAFRNGLALPRLHQLCCRMPTLRRLRVFVLEALGPMALMTFTLFYTFVG 297
Dh 272 SSGC-----ARLPF---FQGLIGALKDL----PKPMLILLVYALNMWIAFPLLPDFTDMG 321
Qy 298 EGLVGVRAEPGEARHHDGYRMSGSLGLFLQCAISLVFSLYMDRLVQRFQ--TRAVY 355
Dh 322 KEVVGGT-----VGEGLYDQGHAGALGIMINSVYLGWMSLIEGLARVVGAKRLMG 375
Qy 356 LASVAAPVVAAGATCLSHVAVWASALRGFTFSALQILPYTLASLYHREKQVFLKRY 415
Dh 376 IVNITL-----ACLAMTV-LVTKSAB-----HFRUSHIM----- 405
Qy 416 GDTGASSESDSLMTSFLPGRPGAPFPNGHVAGAGSGCLPPPPALCGASACDVSRYVYG 475
Dh 466 -----GSGNPPPPA--GVKGGALIAFVAVLG 429
Qy 476 EPTEARV-VP-----GRGICLDLAILDSAFLLSOV-----AP 506
Dh 430 IPLATFSPFIPALASIFGSAGSGGLSLGVLNLAIVPQMFVSVTSQPMDFMGGNLP 489
Qy 507 SLFMGSIYOLQSQVAY 523
Dh 490 AFVVGAVATASAVLSF 506

RESULT 4
YD74_SYNV3
ID YD74_SYNV3 STANDARD: PRT: 544 AA.
AC P74168:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical symporter SLI1374.
GN SLI1374.
OS Synecchocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecchocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima T., Hirosewa M., Sugiyura M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okunura S.,
RA Shampo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabeta S.;
RT Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecchocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.*;
RL DNA Res. 3:109-136(1996).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -I- SIMILARITY: BELONGS TO THE SODIUM:GALACTOSIDE SYMPORTER FAMILY
CC (SGP).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D90912; BAA18257.1; -
CC InterPro: IPR001927; Na_galact_symp.

```

| Query Match | Best Local Similarity | Score | DB 1: | Length |
|---|------------------------------------|--------------------------|---------------------------|--------------------------|
| Matches 105; Conservative 60; Mismatches 179; Indels 126; Gaps 20; | 5.0% 22.3% Pred. No. 0.0058; | 142; 22.3% 0.0058; | DB 1: 22.3% 0.0058; | 544; 22.3% 0.0058; |
| 16 AQLLVNLTLELEVCVLAAG-----ITVYPLILEY-GVEEKFMTWVIGPYL 63 | 105 | 142 | DB 1 | 544 |
| 2 SSSLAEKILHFTTKATYAGDGFPAITANILVFLFTLTDVAGIPALAGSVLMIGKIF 61 | 105 | 142 | DB 1 | 544 |
| 64 GLVCPVLGSADHWGRGRPRPTWALSIGLILSLFIPRAGWLAGLCPDPRLELA 123 | 105 | 142 | DB 1 | 544 |
| 62 DAINDPIIGLSDRTRSRWGRRLPMM---LGMIPALFYTAQWILPHFSDRLTNQMG 117 | 105 | 142 | DB 1 | 544 |
| 124 LILIGVLDFECGVCFT---PLEALLSDFRDPDHCQAVSYAFMISLGGCIGLYLP 179 | 105 | 142 | DB 1 | 544 |
| 118 LFIYVAIA-MAFNICYTVNLPYALPELQYNN-RTRLNSRFPAFISIGSILSLYL 175 | 105 | 142 | DB 1 | 544 |
| 180 AIDNMTSALAPFLVGOECLFGLL-TLFLFVQANTLLVAEAAIGPPEAGLSAPLS 238 | 105 | 142 | DB 1 | 544 |
| 176 YI----LIAAGLPPRPOQFGLGVMISVLSIALMSALRLQEKPE---ILSPSL- 226 | 105 | 142 | DB 1 | 544 |
| 239 PHCCPCRAFLA-----FRMIGAL----- 256 | 105 | 142 | DB 1 | 544 |
| 227 -----RRRLAPLMAAGITILLILAIKSFNLGGSGFYISFILLGLINGGFGFTLR 280 | 105 | 142 | DB 1 | 544 |
| 257 -----LPRHLQICMPPTLRR-LFV--AEICSMALMTFTLE 291 | 105 | 142 | DB 1 | 544 |
| 281 DSAVEBHQLKLENSPSPGVTEMLPLKQL--KIAFSNRAFLFVIGIYLCSALVOLTASI 338 | 105 | 142 | DB 1 | 544 |
| 292 YDFVGEGLYGVPRAPRPTTEARRKHDESVRNGSLGLFLQCAISLVFLSYMRVLVQRFST 351 | 105 | 142 | DB 1 | 544 |
| 339 LVYFVVS--WMLINDQOST-----TALAIVO-CTALVMEFLWQALAOFLDK 381 | 105 | 142 | DB 1 | 544 |
| 352 RAVY-LASVAAPVAAGATCLS-HSAVAVTASAAALGFFFSALQILPYTL 399 | 105 | 142 | DB 1 | 544 |
| 382 KVIYPLGSMVMGAEAGLMLVOPGOVALLYTLAIFAGVSAVAILPMNM 431 | 105 | 142 | DB 1 | 544 |

RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RX MEDLINE=21145593; PubMed=11247674;
 RA McVie-Wylie A.J., Lamson D.R., Chen Y.T.;
 RT "Molecular cloning of a novel member of the GLUT family of
 RT transporters, SLC2A10 (GLUT10), localized on chromosome 20q13.1: a
 RT candidate gene for NIDDM susceptibility";
 RL Genomics 72:113-117(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Stavrides G.S., Hashim Y., Huckle E.J., Deloukas P.;
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Posey S.C., Mihic S.J., Craddock A.L., Mychaleckyj J.C., Dawson P.A.,
 RA Bowden D.W.;
 RT "GLUT10: a novel glucose transporter in the type 2 diabetes linked
 RT region of chromosome 20q12-13.1";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W., Butler A.P., Carder C., Carter N.P., Clee C.M.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Coller R.E., Connor R., Cobby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhani P., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graham D.V., Griffiths C., Griffiths M.N.D., Gilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Leharasaino M.H., Leversha M., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McComachie L.J., McIay K., McMurray A.A.,
 RA Milne S., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Senra H.K., Snowken R., Sims S.,
 RA Skue C.D., Smith M.L., Soderlund C., Steward C.A., Sultson J.E.,
 RA Swann M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitaker P., Willey D.L., Williams L., Williams S.A.,
 RA Winling L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20";
 RL Nature 414:865-871(2001).
 CC -1- FUNCTION: Facilitative glucose transporter (By similarity).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: Widely expressed; highest levels in liver and
 CC pancreas.
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE
 CC TRANSPORTERS SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF321240; AKK26294.1; -
 CC EMBL: AL137188; CAB69822.2; -
 CC EMBL: AF248053; CAB31911.1; -
 CC EMBL: AL031055; CAA19926.2; -
 CC MIM: 606145; -
 DR InterPro: IPR003662; sub_transporter.
 DR InterPro: IPR003663; Sugar_transporter.
 DR Pfam: PF00083; sugar_tr.1.
 DR PRINTS: PR00171; SUGRTSPTSPT.
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 2.

DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE NEG.
 KW Transport: Sugar transporter; Transmembrane; Glycoprotein;
 KM Multigene family.
 FT DOMAIN 1 15
 FT TRANSMEM 16 36
 FT DOMAIN 37 48
 FT TRANSMEM 49 69
 FT DOMAIN 70 77
 FT TRANSMEM 78 98
 FT DOMAIN 99 106
 FT TRANSMEM 107 127
 FT DOMAIN 128 133
 FT TRANSMEM 135 155
 FT DOMAIN 156 166
 FT TRANSMEM 167 187
 FT DOMAIN 188 233
 FT TRANSMEM 234 254
 FT TRANSMEM 255 269
 FT TRANSMEM 270 290
 FT DOMAIN 291 298
 FT TRANSMEM 299 319
 FT DOMAIN 320 414
 FT TRANSMEM 415 435
 FT DOMAIN 436 445
 FT TRANSMEM 446 466
 FT DOMAIN 467 476
 FT TRANSMEM 477 497
 FT DOMAIN 498 541
 FT CARBOHYD 334 334
 SQ SEQUENCE 541 AA: 56911 MW: 60644525FA136908 CRC64;

Query Match 4.8%; Score 136; DB 1; Length 541;
 Best Local Similarity 21.9%; Pred. No. 0.015;
 Matches 130; Conservative 72; Mismatches 210; Indels 182; Gaps 27;
 16 AOLLVNLTFEGVCLAGITVPPLELVG---VEEKFMVGLGIPVLGVCPVLG 72
 12 ASVSLGGLTFGLVAVISGA--LPLQDLFGISCEQEFVLSLGLALLSVGFLI 69
 73 SASDHWGRVGRRRPRIMALSIGILSLPLPRAGLACDPPLLELILVGLL 132
 70 DC-----YGRQALIGS-NLVLAGSLTLAG-----SLAMVLGRAVY 108
 133 DPCGVCFPLLEALSDFRDPDHCQANSVAFMISLC---GCLGYLLPAIDMTS-- 186
 109 GFAISLSMACTIVSELY-GPRQGVLSLYEAGITVGLISYALNVALAGTPMGRHM 167
 187 ---ALAPYLGTOBECFGLTLTFLCVAAATLVAEALGPTEPAGLSAPLSHCP 243
 168 FGNATAP-----AVLOSLSLFLP--AGT---DETATNKKDILPLQGEAPRLGFG-- 212
 244 CRRALAFRNIGALLPRLHOLCCMRPTRLRLFAVELCSMMALMTFTLFYDFEGELYOG 303
 213 -RRRYFELDLFRANDMRG---RTVGLGLVLEQQLTGQPNVLCVA--STIESVGFHG 266
 304 VPRAEFGTEARRHYDEGVMSGLGLFLOCAISLVSWDRLYQRTGTAUVASYA--- 360
 267 -----SSAVLASVGL--CAVKAATATLTMGLVDRGRRLALLGACALMA 308
 361 -----AFPVAGATCUSHSAVAVTASALNG---FFFSALQILPYTLASTLYHR 405
 309 LSVSGIGLVSPAVPMDSGSCL--AVPNATGQTGLGDSGLDSSLPPIPT-----N 360
 406 EKQVFLPKYRGDTGASSDSLMTSFLPGKPGAPRPNGHVAGSGGLRPPPALGASA 465
 361 EDQ-----REPLISTAKTKRPHRSGDPSAPRDLALSALPGP----- 399
 466 CDVSVAWVVGEPLEAVVVGSG-----ICDLALIDSAF-----LISQVAP-- 506
 400 -----LPARGHALLRWTAALCLAMFVASFSGFGPVTWLVISLSELYPVE 442
 507 -----SLFMG-----SIVQLSGSVTVAVMSAA-GLGLVAVIY 536

```

DB 443 IRGRAFAFCNSFWMAANLFTSLFLDITGLTGLSWFLYLGLTAVGLGTYLF 496
RESULT 6
TCRL_ECOLI STANDARD; PRT; 399 AA.
ID TCRL_ECOLI
AC P02982;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Tetracycline resistance protein, class A (TETRA(A)).
GN TETRA.
OS Escherichia coli.
OC Plasmid Rpl.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC TRANSPOSON-Tn1721;
RX MEDLINE=92192465; Pubmed=1312499;
RA Allmeier H., Cresnar B., Greck M., Schmitt R.;
RT "Complete nucleotide sequence of Tn1721: gene organization and a
RT novel gene product with features of a chemotaxis protein."
RL Gene 111:11-20(1992).
RN [2]
RP SEQUENCE OF 85-399 FROM N.A.
RC STRAIN=DH1;
RA Fletcher J.N., Hart C.A., Batt R.M., Saunders J.R.;
RN Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC PLASMID-RPL; TRANSPOSON-Tn1721;
RX MEDLINE=83292270; Pubmed=6310527;
RA Waters S.H., Rogowsky P., Ginsted J., Altenbuchner J., Schmitt R.;
RT "The tetracycline resistance determinants of Rpl and Tn1721:
RT nucleotide sequence analysis."
RL Nucleic Acids Res. 11:6089-6105(1983).
RN [4]
RP TOPOLOGY.
RX MEDLINE=92388137; Pubmed=1517220;
RA Allard J.D., Bertrand K.P.;
RT "Membrane topology of the pBR322 tetracycline resistance protein.
RT TETRA-PhoA gene fusions and implications for the mechanism of TetA
RT membrane insertion."
RL J. Biol. Chem. 267:17809-17819(1992).
CC -1- FUNCTION. RESISTANCE TO TETRACYCLINE BY AN ACTIVE TETRACYCLINE
CC EFFLUX. THIS IS AN ENERGY-DEPENDENT PROCESS THAT DECREASES THE
CC ACCUMULATION OF THE ANTIBIOTIC IN WHOLE CELLS. THIS PROTEIN
CC FUNCTIONS AS A METAL-TETRACYCLINE/H+ ANTI-PORTER.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSDUCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X61367; CAA3643.1; -
DR EMBL: L29404; AAB83545.1; -
DR EMBL: X00006; CAA24909.1; -
DR PIR: A03509; YTECK1.
DR InterPro: IPR001958; TCR_TetA.
DR InterPro: IPR003662; sub_transporter.
DR Pfam: PF00083; sugat tr; 1.
DR PRINTS: PR01035; TCR_TETRA.
DR PROSITE: PS00216; SUGAR_TRANSPORT_1; UNKNOWN_1.
KW Antibiotic resistance; Transmembrane; Inner membrane; Transport;

```

```

KW Symport; Transposable element; Plasmid.
FT DOMAIN 1 7
FT TRANSMEM 8 27
FT DOMAIN 28 45
FT TRANSMEM 46 66
FT DOMAIN 67 79
FT TRANSMEM 80 100
FT DOMAIN 101 103
FT TRANSMEM 104 124
FT DOMAIN 125 138
FT TRANSMEM 139 159
FT DOMAIN 160 160
FT TRANSMEM 161 181
FT DOMAIN 182 210
FT TRANSMEM 211 231
FT DOMAIN 232 246
FT TRANSMEM 247 267
FT DOMAIN 268 277
FT TRANSMEM 278 298
FT DOMAIN 299 299
FT TRANSMEM 300 320
FT DOMAIN 321 339
FT TRANSMEM 340 360
FT DOMAIN 361 364
FT TRANSMEM 365 385
FT DOMAIN 386 399
FT CONFLICT 5 5
FT CONFLICT 55 55
FT CONFLICT 75 75
FT CONFLICT 84 84
FT CONFLICT 201 203
SQ SEQUENCE 399 AA; 42240 MW; 2984278B5478374 CRC64;

```

```

Query Match 4.4%; Score 125.5; DB 1; Length 399;
Best Local Similarity 22.3%; Pred. No. 0.062;
Matches 104; Conservative 57; Mismatches 136; Indels 169; Gaps 27;

```

```

OY 19 LVLNLLTFGLVCAAGTTPPL--LLEVGVEKEFT-----AVLGPGVLGVCPPL 71
DB 7 LIVLITFVALD--AVGIGLIMPVLPGLRLDLVSHNDVTAHYGLLALVALMOFRCAVPL 63
OY 72 GSASDHMRGRYGR--PPFIALSLGILSLFLPRAGWLAGL 112
DB 64 GALSU---RGRRPVLLVSLAGAAVDYAINATPFLWLYIGRVA----- 106
OY 113 LCPFRPLEALLILVGLIDFCGVCPTPEALLSLDFRDPDCHRCQAVSYVAFMISLGG 172
DB 107 -----GITGATGAVA-----GAVYADITDGERARH---FGFM---SA 138
OY 173 CLGYLLPAIDMTSALAPYLGTOEBCLEGLTLFLTCVAATLVAEAAALGPPEPAEGL 232
DB 139 CFGFGMVA-----GVLG---GLM----- 154
OY 233 SAPSLPHCCPCORARLAPRNG-----ALPRLHOLCCMRPTLRLEVAELCS--W--- 282
DB 155 -GGGSPH-APFPAALANGLNLFTGCLLPESHK---GERRPLREKLNPLASRRMARG 208
OY 283 ---MALMTFTLYTDVGEGLVGVPRRA--EPGTERRHYDE---GVBMGSLGLEFLQCAI 334
DB 209 MIVVAALAAVFFIMQVQO-----VPALMWVTFGDRRHMDATTTGISLAFAGLHSLAQ 263
OY 335 SLVSLVMDRLVQRFGRTAIVLASVA---APVVAAGAT---CLSHSVAVVTSAALTGTF 389
DB 264 AMITGPVAARLGER---PALMIGMIDTGTGTLFAFRGWMAFPIWLLASG---GIGM 317
OY 390 SALQTLPTTLASLVHREKOVFLPKRYRGDTGGASSSDSLTSPFLPGP 435
DB 318 PALQAM---LSRQVDEERO-----GQLQSLAALTSLTSLIV-GP 352
RESULT 7
PUR8_STRLP

```

ID PUR8_STRLP STANDARD; PRT; 503 AA.
AC P42670;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Puromycin resistance protein pur8.
GN PUR8.
OS Streptomyces lipmanii (Streptomyces alboniger).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID:1892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 12461;
RX MEDLINE=94109397; PubMed=7916693;
RA Terceiro J.A., Laccalle R.A., Jimenez A.;
RT "The pur8 gene from the pur cluster of streptomyces alboniger encodes
RT a highly hydrophobic polypeptide which confers resistance to
RT puromycin.";
RL Eur. J. Biochem. 218:963-971(1993).
CC -!- FUNCTION: MAY BE INVOLVED IN ACTIVE PUROMYCIN EFFLUX ENERGIZED BY
CC A PROTON-DEPENDENT ELECTROCHEMICAL GRADIENT. IN ADDITION, IT COULD
CC BE IMPLICATED IN SECRETING N-ACETYLPUROMYCIN, THE LAST
CC INTERMEDIATE OF THE PUROMYCIN BIOSYNTHESIS PATHWAY, TO THE
CC ENVIRONMENT.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X76855; CAAS4186.1; -
DR Pfam: PF00083; sugar_tr.1; -
KW Antibiotic resistance; Antibiotic biosynthesis; Transmembrane;
KW Transport.
FT DOMAIN 1 24 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 25 45 POTENTIAL.
FT DOMAIN 46 64 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 65 85 POTENTIAL.
FT DOMAIN 86 92 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 93 113 POTENTIAL.
FT DOMAIN 114 122 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 123 143 POTENTIAL.
FT DOMAIN 144 152 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 153 173 POTENTIAL.
FT DOMAIN 174 181 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 182 202 POTENTIAL.
FT DOMAIN 203 212 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 213 233 POTENTIAL.
FT DOMAIN 234 241 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 242 262 POTENTIAL.
FT DOMAIN 263 280 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 281 301 POTENTIAL.
FT DOMAIN 302 313 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 314 334 POTENTIAL.
FT DOMAIN 335 346 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 347 367 POTENTIAL.
FT DOMAIN 368 371 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 372 392 POTENTIAL.
FT DOMAIN 393 422 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 423 443 POTENTIAL.
FT DOMAIN 444 461 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 462 482 POTENTIAL.
FT DOMAIN 483 503 CYTOPLASMIC (POTENTIAL).
SO SEQUENCE 503 AA; 51852 MW; FA3B814DF9A3EB74 CRC64;

Query Match 4.2%; Score 121.5; DB 1; Length 503;
Best Local Similarity 23.2%; Pred. NO. 0.15;
Matches 131; Conservative 61; Mismatches 175; Indels 197; Gaps 32;
OY 26 FGLEVCCLAQITVY-----PPLLEVGVEEKFMITV-----LGIGPVGLVCYPL 70
DB 26 WGLVYLAAQLLVLDGTFVYNIALPSVGRDLGMSDPSRQWVITAYTLAAGLL-----L 79
OY 71 IGSASDHRGR-----YGRRRPTMALSGIILSLFLPAGN-LAGLL--CPDRPLEL 122
DB 80 LG-----GRVADAFGRRRIF-----AVGIL-----GGLASLLGGAAPDGTFL 119
OY 123 ALLITGVLLDFCGVCFPLLEALLSDLEPDDHCRQASVAFMISLGGCGYLLPAID 182
DB 120 AVALGCV-----FAALLAPAL-ALINTLTETEGEGRKAFGYGAVSGCAVGLLAG--- 171
OY 183 WDSALAPYLTQEBCLFGLT-LIFLCVATLLVAEEALGPTPEAGLSAPSLSPH 240
DB 172 -----GLTFEYLDWRMCCLVYNAPVALLALG----- 197
OY 241 CCPCARLAFRNLGALLPRLHOLCCMRPTLRFLVAEL--CSMALMTFTLTVDVGE 298
DB 198 --CR-----LLPRDR-----RTGRVRLDLPGLTGGCGVAIVYAF----- 232
OY 299 GLYGVPAEPCTEARRHVDEG--VRMGSGLFLCALSLVPSLVMDRLVQRFTRAVYL 356
DB 233 -----AEBSGSGDPLVYRLVLGVLMVAFLVRRQDPLP----- 271
OY 357 ASVAAPVAAGATCLSHSVAVTASALTG-----FTFSALQILPYT--LASLYHREKQVF 410
DB 272 PCVVAHRVAGS-----FLVVGLPQIGLFLPLFYLYLOGILDYSPVLTV-----AF 320
OY 411 LPKYGDGTGGASEDSLMSTFLRGPX-----GAFPPKNVAGSGSGLL-----PPPALC 461
DB 321 LPLGLGIAVSS--LIAARLLPRTLRPTLVGALL--AAAGMALTLRLPPTP-- 370
OY 462 GASACDVSVRVVGPTEARVVRGRI-CLDLAILDSAPLSQVAP-----SLPMGSIV 514
DB 371 -----QVYLTHLLPRAQLILGIGICGMMKPAHHTA--TARVAPHEGAAAVVNSAQ 419
OY 515 QLSQSVTAIVMSAAGLGLVAIFYA 538
DB 420 QVGALGVALLVNTVSTGATAVLYA 443
RESULT 8
AMPG_ECOLI STANDARD; PRT; 491 AA.
AC P36670;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE AmpG protein.
GN AMPG OR B0433 OR Z0536 OR ECS0487.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94049112; PubMed=8231804;
RA Lindquist S., Weston-Hafer K., Schmidt H., Pul C., Korfmann G.,
RA Erickson J., Sanders C., Martin H.H., Normark S.;
RT "AmpG, a signal transducer in chromosomal beta-lactamase induction.";
RL Mol. Microbiol. 9:703-715(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-SN0301-1, SN0301-3, AND SN0301-5;
RX MEDLINE=95291453; PubMed=7773404;
RA Schmidt H., Korfmann G., Barth H., Martin H.H.;
RT "The signal transducer encoded by ampG is essential for induction of
RT chromosomal AmpC beta-lactamase in Escherichia coli by beta-lactam

RT antibiotics and 'unspecific' inducers." ;
 RL Microbiology 141:1085-1092(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / MG1655;
 RX MEDLINE=97426617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. II, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Valdes J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y. ;
 RL "The complete genome sequence of *Escherichia coli* K-12." ;
 RL Science 277:1453-1474(1997).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC Robert's D., Allen E., Araujo R., Aparicio A., Chung E., Davis K.,
 RA Duncan M., Federspiel N., Hyman R., Kalman S., Komp C., Kudr O.,
 RA Lew H., Lin D., Namath A., Oefner P., Schramm S., Davis R.W. ;
 RL submitted (JAN-1997) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
 RX MEDLINE=21074935; PubMed=11206551;
 RA Perna N.T., Plunkett G. II, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Postal G., Heckett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lin A., Dimalanta E.T., Potamousis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R. ;
 RL "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7." ;
 RL Nature 409:529-533(2001).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN-O157:H7 / RIMD 0509952;
 RX MEDLINE=21156231; PubMed=11258796;
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,
 RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
 RA Kuhnra S., Shiba T., Hattori M., Shinagawa H.,
 RA "Complete genome sequence of enterohemorrhagic *Escherichia coli*
 RT O157:H7 and genomic comparison with a laboratory strain K-12." ;
 RL DNA Res. 8:11-22(2001).
 CC -1- FUNCTION: PROBABLY ACTS AS A PERMEASE IN THE BETA-LACTAMASE
 CC INDUCTION SYSTEM AND IN PEPTIDOGLYCAN RECYCLING.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 CC (Probable).
 CC -1- SIMILARITY: TO H.INFLUENZAE HI0350 AND TO YEAST YBR220C.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL, S67816; AAB28884.1; -;
 DR EMBL, X82158; CAA57651.1; -;
 DR EMBL, X82159; CAA57652.1; -;
 DR EMBL, X82160; CAA57653.1; -;
 DR EMBL, AE000149; AAC73536.1; -;
 DR EMBL, U82664; AAB40189.1; -;
 DR EMBL, AE005222; AAG54783.1; -;
 DR EMBL, AP002551; BAB31910.1; -;
 DR PIR, S37391; S37391.
 DR Ecocore: EG12183; amp.
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Transprot; Transmembrane; Inner membrane; Complete proteome.
 FT TRANSMEM 13 33 POTENTIAL.
 FT TRANSMEM 81 105 POTENTIAL.
 FT TRANSMEM 110 128 POTENTIAL.
 FT TRANSMEM 174 197 POTENTIAL.
 FT TRANSMEM 224 244 POTENTIAL.

| Query Match | 4.2%: Score 119.5; DB 1; Length 491; |
|---|--------------------------------------|
| Best Local Similarity 21.0%: Pred. No. 0.2; Mismatches 126; Conservative 76; Indels 193; Gaps 30; | |
| 9 RLIRRRKQALLVNLITFGLEYCLAAGITVYPRLLLEVGVEEKFMT-----VLGIGPV 62 | |
| 7 RIFQQRRAIILTI-----LGFASGL-----PLALITSGTLDAMMTVEINDIKTIGFSL 54 | |
| 63 LGLVGV-----PLIGSASDHWKRGYRRRPITWALSIGILSL-----FLIP--RAGMIA 110 | |
| 55 VGOAVVFKEFLSPMDRYTPPF---FGRRRGWLLTQIILLVAIAAMGFLPGETOLDRWMA 111 | |
| 111 GLICDPRPRLALLILGVGLDFGQVCFPLEALLSDLPFRDPH--CQOAVSYAF--- 166 | |
| 112 ALAV-----VIAFCSASODIVEDMKTDVDPRAEEGAGAAISVAGIRLG 155 | |
| 167 MISLGCGYLLPAIDWPTSAIAPLIGTOEELFGILLTLEFCVAATLLVAEEALGPT 226 | |
| 156 MLVSGGLALML--ADKW-----LGMQG--MYMLAMAILPCLITATLAP-----PT 198 | |
| 227 EPAEGLSAPLSPHCCPCRARLAFRNGLALLPRHLQCLCRMDRTLRLFEVLE----- 279 | |
| 199 D-----TIP-----VPTLEQAVVAPLDRDFEGHN 222 | |
| 280 CSMWALMTFTL-----FYTDVGEGLVGVRAERGTFRARNHYDEGVMSGLGLEQCA 333 | |
| 223 NAMILLILLIVLKIGDAFAMSITTFELIRGV-----GPDAG--EVGVNKKTLGTLATTIV 274 | |
| 334 ISLVESLVMDRLVQRFGRFVAVLASAAPPAAGATCLSHSAVVYATASALTGFPELSAQ 393 | |
| 275 GALYGIIMQRIQL--FRALLIFGI-----LOGASNAGTW 307 | |
| 394 ILPYTILASIVHREKQVFLPKYRGIDTGASSSEDSMLTSPFLRGPGRGAPFPNGHVGAGSGGI 453 | |
| 308 ILITIDKHLYSKGAVFPEFNLCGKGTSALFVALLMT-----LCNKSPTSATQFAL 356 | |
| 454 LPPPALCGASACDYSVRVYVGEPIEARVYPPRGICLDLAILDSAPILSOVA---PSLEM 510 | |
| 357 L-----SALSASVGRYVG--PAAGWFEAHGW-----STFYLESVAAYVAGLL 398 | |
| 511 GSTVQLSQ-----SVTAVMYS-----AAGLGLVAVIYFATQVVFDPKSLDAKYS 552 | |
| 399 LIAVQRTLEIYTRVNDNFISRTAIPACIYAFAMTTLAAGVSLAVLWLLL-LTMDALDLTHFS 457 | |
| RESULT 9 | |
| YHJF_ECOLI | |
| ID YHJF_ECOLI | STANDARD: PRT: 440 AA. |
| AC P37643; | |
| DT 01-OCT-1994 (Rel. 30, Created) | |
| DT 01-OCT-1994 (Rel. 30, Last sequence update) | |
| DT 01-MAR-2002 (Rel. 41, Last annotation update) | |
| DE Hypothetical metabolite transport protein yhjE. | |
| GN YHJF OR B3523. | |
| OS Escherichia coli. | |
| OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | |
| OC Escherichia. | |
| OX NCBI_TaxID=562; | |
| RN [1] | |

| Accession | Protein Name | Length (aa) | Weight (kDa) | PI (pI) | Source | Notes |
|-----------|--|-------------|--------------|---------|--------|-------|
| RP | SEQUENCE FROM N.A. | | | | | |
| RC | STRAIN-K12 / MG1655: | | | | | |
| RX | MEDLINE-94316500; PubMed-8041620; | | | | | |
| RA | Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.; | | | | | |
| RT | "Analysis of the Escherichia coli genome. V. DNA sequence of the | | | | | |
| RT | region from 76.0 to 81.5 minutes." | | | | | |
| RL | Nucleic Acids Res. 22:2576-2586(1994). | | | | | |
| CC | -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane | | | | | |
| CC | (potential). | | | | | |
| CC | -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. STRONG, TO | | | | | |
| CC | H. INFLUENZAE HI0281 AND HI0418. | | | | | |
| CC | ----- | | | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration with | | | | | |
| CC | the Swiss Institute of Bioinformatics and the EMBL outstation - | | | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on | | | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | | | |
| CC | entities requires a license agreement (See http://www.isb-slb.ch/announce/ | | | | | |
| CC | or send an email to license@isb-slb.ch). | | | | | |
| CC | ----- | | | | | |
| DR | EMBL; U00039; AAB18499.1; -. | | | | | |
| DR | EMBL; AE000429; AAC76548.1; -. | | | | | |
| DR | EcoGene; EGI12249; yJbE. | | | | | |
| DR | InterPro; IPR003662; sub.transporter. | | | | | |
| DR | Pfam; PF00083; sugar_tr; 1. | | | | | |
| DR | PROSITE; PS00216; SUGAR_TRANSPORT_1; 1. | | | | | |
| DR | PROSITE; PS00211; SUGAR_TRANSPORT_2; FALSE_NEG. | | | | | |
| KW | Hydrophobic protein; Transport; Transmembrane; Inner membrane; | | | | | |
| KW | Complete proteome. | | | | | |
| FT | DOMAIN 1 34 | | | | | |
| FT | TRANSMEM 35 55 | | | | | |
| FT | DOMAIN 56 66 | | | | | |
| FT | TRANSMEM 67 87 | | | | | |
| FT | DOMAIN 88 108 | | | | | |
| FT | TRANSMEM 109 129 | | | | | |
| FT | TRANSMEM 130 150 | | | | | |
| FT | DOMAIN 151 167 | | | | | |
| FT | TRANSMEM 168 188 | | | | | |
| FT | DOMAIN 189 192 | | | | | |
| FT | TRANSMEM 193 213 | | | | | |
| FT | TRANSMEM 214 248 | | | | | |
| FT | TRANSMEM 249 269 | | | | | |
| FT | TRANSMEM 270 289 | | | | | |
| FT | DOMAIN 290 310 | | | | | |
| FT | TRANSMEM 311 320 | | | | | |
| FT | TRANSMEM 321 341 | | | | | |
| FT | TRANSMEM 342 345 | | | | | |
| FT | TRANSMEM 346 366 | | | | | |
| FT | DOMAIN 367 384 | | | | | |
| FT | TRANSMEM 385 405 | | | | | |
| FT | DOMAIN 406 410 | | | | | |
| FT | TRANSMEM 411 431 | | | | | |
| FT | DOMAIN 432 440 | | | | | |
| SO | SEQUENCE 440 AA; 47208 MW; 36233029ADCEB5B CQC64; | | | | | |

| | | | | |
|-----------------------|--------------|-----------------|----------------|-------------|
| Query Match | 4.18; | Score 117; | DB 1; | Length 440; |
| Best Local Similarity | 28.48; | Pred. No. 0.27; | | |
| Matches 48; | Conservative | 29; | Mismatches 54; | Indels 38; |
| | | | Gaps | 10. |

| | | | |
|----|-----|--|-----|
| QY | 53 | MTWVLGICPVILGCVPLIGSASDHMRGRYGRPRPFIALSGILLSLFLI-PRAGWLAG | 111 |
| | | : : | : : |
| Db | 295 | MAAVIGFC-----VMVPAGLLAD---AFGRKRSMTITITLILFALFAFNPLG---- | 341 |
| | | : : | : : |
| QY | 112 | LICDPRR-LELAILIIGVGLDFPGGYCFPFLLESLDLPDPDPCHQAYSVAFMISL | 170 |
| | | : : | : : |
| Db | 342 | -----SGNFIIVAFALFLPLSLMG----LTFGGMGLLPFLF--PFEVR--YTGASFSTYV | 389 |
| | | : : | : : |
| QY | 171 | GGCIGLYLPAIDMDTSALAPYLGTOECLFGILTL-IFLTCVAAALILVA | 218 |
| | | : : | : : |
| Db | 390 | ASILG-----ASVAPITAMLIQNTYIGGAGVGLTAAAGLITLLA | 428 |
| | | : : | : : |

| RESULT | 10 | | | |
|--------|--|-----------|------|--------------------------|
| ID | SMVA_SALTY | STANDARD: | PRT: | 495 AA. |
| AC | P37594. | | | |
| DT | 01-OCT-1994 (Rel. 30, Created) | | | |
| DT | 01-MAR-2002 (Rel. 41, Last sequence update) | | | |
| DT | 01-MAR-2002 (Rel. 41, Last annotation update) | | | |
| DE | Methyl viologen resistance protein smva. | | | |
| GN | SMVA OR STM1574. | | | |
| OS | Salmonella typhimurium. | | | |
| OC | Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; | | | |
| OC | Salmonella. | | | |
| OX | NCBI_TaxID=602; | | | |
| RP | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=SL1303; | | | |
| RC | MEDLINE=95011654; PubMed=7926834; | | | |
| RA | Hongo E., Moriyama M., Mita K., Machida I., Hama-Inaba H., Tsuji H., | | | |
| RA | Ichimura S., Noda Y.; | | | |
| RA | "the methyl viologen-resistance-encoding gene smva of Salmonella | | | |
| RA | typhimurium."; | | | |
| RL | Gene 148:173-174(1994). | | | |
| RL | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=LT2 / SGSC1412 / ATCC 700720; | | | |
| RC | MEDLINE=21534948; PubMed=11677609; | | | |
| RA | McClelland M., Sanderson K.E., Speith J., Clifton S.W., Latreille P., | | | |
| RA | Courtney L., Potyollik S., Ali J., Dante M., Du F., Hou S., Layman D., | | | |
| RA | Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E., | | | |
| RA | Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M., | | | |
| RA | Waterston R., Wilson R.K.; | | | |
| RA | "Complete genome sequence of Salmonella enterica serovar Typhimurium | | | |
| RA | LT2."; | | | |
| RL | Nature 413:852-856(2001). | | | |
| CC | -1 SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane | | | |
| CC | (Potential). | | | |
| CC | -1 SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN | | | |
| CC | AS THE DRUG RESISTANCE TRANSLOCASE FAMILY). | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/annouce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL: D26057; BAA05055.1; -. | | | |
| DR | EMBL: AE008769; AAL20492.1; -. | | | |
| DR | StyGene: SG10384; smva. | | | |
| KW | Transmembrane; Inner membrane; Complete proteome. | | | |
| FT | TRANSMEM 5 25 | | | POTENTIAL. |
| FT | TRANSMEM 44 64 | | | POTENTIAL. |
| FT | TRANSMEM 73 93 | | | POTENTIAL. |
| FT | TRANSMEM 96 116 | | | POTENTIAL. |
| FT | TRANSMEM 135 135 | | | POTENTIAL. |
| FT | TRANSMEM 158 178 | | | POTENTIAL. |
| FT | TRANSMEM 192 212 | | | POTENTIAL. |
| FT | TRANSMEM 220 240 | | | POTENTIAL. |
| FT | TRANSMEM 260 280 | | | POTENTIAL. |
| FT | TRANSMEM 299 319 | | | POTENTIAL. |
| FT | TRANSMEM 327 347 | | | POTENTIAL. |
| FT | TRANSMEM 357 377 | | | POTENTIAL. |
| FT | TRANSMEM 391 411 | | | POTENTIAL. |
| FT | TRANSMEM 469 489 | | | POTENTIAL. |
| FT | TRANSMEM 144 144 | | | A -> R (IN REF. 1). |
| FT | CONFLICT 147 147 | | | G -> V (IN REF. 1). |
| FT | CONFLICT 182 182 | | | V -> D (IN REF. 1). |
| FT | CONFLICT 198 198 | | | A -> V (IN REF. 1). |
| FT | CONFLICT 227 227 | | | F -> Y (IN REF. 1). |
| FT | CONFLICT 447 447 | | | L -> Y (IN REF. 1). |
| FT | CONFLICT 454 457 | | | ALD -> GRLT (IN REF. 1). |
| SO | SEQUENCE 495 AA; 52134 MW; 961F77C748C0A164 CRC64; | | | |

OY 377 --VTASALTGFTFSALQILPYT--IASLYHREKQVFLPKYRGDNGCASSEDSLMTFLP 433
 DB 459 PLIVAPMVRRAATPLPAQAVKFTGLGAV-----VRLPAMG-----SLIP 499
 OY 434 GPKGAPFPNGHVGAGSGSLP-----PPPA-----LCGASACDVSRYVVGPEPTEAR 481
 DB 500 GYIAAAYL-----AAALAVAVLARWFRFRPARAPARLPLMACGA--DLTVRMQYTTATSPAE 553
 OY 482 VYPRGICLDLALDSAFLLSOVAPSLFMGSTVOLSQV 520
 DB 554 --PLQRFEGDVLRPDITIEVHTAESRYMAERITYRTAV 590
 RESULT 12
 PHDK_NOC SK STANDARD: PRT: 473 AA.
 ID PHDK_NOC SK STANDARD: PRT: 473 AA.
 AC 024723;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Probable 1-hydroxy-2-naphthoate transporter.
 OS Nocardioides sp. (strain KP7).
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Propionibacteriaceae; Nocardioides.
 OX NCBI_Taxid=35761;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=97474276; PubMed=9335300;
 RA Iwabuchi T., Harayama S.;
 RT "Biochemical and genetic characterization of 2-carboxybenzaldehyde
 dehydrogenase, an enzyme involved in phenanthrene degradation by
 Nocardioides sp. strain KP7."
 RL J. Bacteriol. 179:6488-6494(1997).
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE TRANSPORT OF 1-HYDROXY-2-
 NAPHTHOATE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (potential).
 CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL: AB000735; BAA23264.1;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr.1
 DR PROSITE: PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE: PS00217; SUGAR_TRANSPORT_2; FALSE_NEG.
 KW Transport; Transmembrane; Inner membrane.
 FT TRANSMEM 1 25 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 26 46 1 (POTENTIAL).
 FT TRANSMEM 47 59 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 60 80 2 (POTENTIAL).
 FT DOMAIN 81 91 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 92 112 3 (POTENTIAL).
 FT DOMAIN 113 121 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 122 142 4 (POTENTIAL).
 FT DOMAIN 143 153 5 (POTENTIAL).
 FT TRANSMEM 154 174 6 (POTENTIAL).
 FT DOMAIN 175 180 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 181 201 6 (POTENTIAL).
 FT TRANSMEM 202 263 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 264 284 7 (POTENTIAL).
 FT DOMAIN 285 304 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 305 325 8 (POTENTIAL).
 FT DOMAIN 326 330 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 331 351 9 (POTENTIAL).
 FT DOMAIN 352 354 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 355 375 10 (POTENTIAL).
 FT DOMAIN 376 398 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 399 418 11 (POTENTIAL).
 FT DOMAIN 419 421 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 422 444 12 (POTENTIAL).
 FT DOMAIN 445 473 CYTOPLASMIC (POTENTIAL).
 SQ SEQUENCE 473 AA; 49109 MW; D6D765D37626D08A CRC64;
 Query Match 3.9%; Score 111; DB 1; Length 473;
 Best Local Similarity 24.3%; Pred. No. 0.78;
 Matches 114; Conservative 66; Mismatches 180; Indels 110; Gaps 27;
 OY 12 RHRKQQLLVNLTFF--GLEVCIAAGITVVPPLLEVEGEKFMVVGIPVLGYCVP 69
 DB 22 RRRITVLMVFLTMVADGMDITLASHL--PPVIRDMGVPSAVTLVSLGVVAAAGL 79
 OY 70 LIGSASDHWGRGRYRRRPRTMALSLGILSLFLPRAG--WLAGLCPDPRPLEALLIIG 128
 DB 80 VSGPVADRW--GRKG-----VTVGVFV--LPCLATAGLGLG---DIHSPALRIISC 125
 OY 129 VGLIDFCGCVCTPLPALLSDLEFRDPDHC--QAYSVAAMISLGCLGLPLAIDMTSA 187
 DB 126 FGL---GAVMPVAL--TIVADWM--PKARRAQMVSIAFAGVGSIIGAYLAA-----A 172
 OY 188 LAPYIGTOECLE--GLLTLIFLTCVATLLVAEEA-----ALGPTEPARG 231
 DB 173 VPTLGMQVWMLTAGLAPLILPFVYA--LVPEPALIISVRKGIPEARRISALALVAPDRD 230
 OY 232 LSAPSLSPCCPCORARLARFNALPRLHOLCCRPRLRLFLVLCSSMALMTFLF 291
 DB 231 IAGVDL-----RAGL--TLGAGEVR-----AKLFAFILLRPLIGVTLIM 270
 OY 292 YIDFVGEG---LYGVP---RAEPTEARRHYDEGVMSGLFLQCAISLVSIVND 343
 DB 271 GVFVVGSGGLLVLYQMPMLAPAPGLST---VESGLIVAMGWS---ALIGOLITTA 322
 OY 344 RIVORGRTRAVYLSVAAPVAAGATCLSHSVAVVAAVSLGFTF-----SAQI-L 395
 DB 323 FIKRP--DFPILAAFTFVS-----VGLLIVAFGTGFGFTLLFALIGSL 371
 OY 396 PYTLASLYHREKQVFLPKYR---GDTGASSEDSLSMTFLPGPKGAPF 441
 DB 372 PNTAAMQSVTTLAYEEFRATGMSAGFAGRGITLTTCALGSLTLAGAF 421
 RESULT 13
 CA1A_HUMAN STANDARD: PRT: 680 AA.
 ID CA1A_HUMAN STANDARD: PRT: 680 AA.
 AC 003692;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OX NCBI_Taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=92109659; PubMed=1764025;
 RA Thomas J.T., Gresswell C.J., Rash B., Nicolai H., Jones T.,
 RA Solomon E., Grant M.E., Boot-Handford R.P.;
 RT "The human collagen X gene. Complete primary translated sequence and
 RT chromosomal localization."
 RT Biochem. J. 280:617-623(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=93012005; PubMed=1397333;
 RA Reichenberger E., Belter F., Luvallie P., Olsen B.R., von der Mark K.,
 RA Bertling W.M.;
 RT "Genomic organization and full-length cDNA sequence of human collagen

RT X.";
 RL FEBS Lett. 311:305-310(1992).
 RL [3]
 RP SEQUENCE FROM N.A.
 RA Beier F., Lammi M.B., von der Mark K.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Williams S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 52-680 FROM N.A.
 RX MEDLINE=92267014; PubMed=1587271;
 RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
 RT "Cloning of the human and mouse type X collagen genes and mapping of
 the mouse type X collagen gene to chromosome 10.";
 RL Eur. J. Biochem. 206:217-224(1992).
 RN [6]
 RP SEQUENCE OF 561-666 FROM N.A.
 RX MEDLINE=91243838; PubMed=2037056;
 RA Apte S., Mattei M.-G., Olsen B.R.;
 RT "Cloning of human alpha 1(X) collagen DNA and localization of the
 COL10A1 gene to the q21-q22 region of human chromosome 6.";
 RL FEBS Lett. 282:393-396(1991).
 RN [7]
 RP SEQUENCE OF 547-655 FROM N.A.
 RX MEDLINE=92077285; PubMed=1743401;
 RA Reichenberger E., Aigner T., von der Mark K., Stoeck H., Berling W.;
 RT "In situ hybridization studies on the expression of type X collagen
 in fetal human cartilage.";
 RL Dev. Biol. 148:562-572(1991).
 RN [8]
 RP REVIEW ON VARIANTS.
 RX MEDLINE=97255959; PubMed=9101290;
 RA Kuwamura H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, III, and XI), fibril-
 associated collagen (type IX), and network-forming collagen (type X)
 cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [9]
 RP VARIANTS SMCD ASP-598 AND PRO-614.
 RX MEDLINE=94136476; PubMed=8304336;
 RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
 RA Grant M.E., Boot-Handford R.P.;
 RT "Amino acid substitutions of conserved residues in the
 carboxyl-terminal domain of the alpha 1(X) chain of type X collagen
 occur in two unrelated families with metaphyseal chondrodysplasia
 type Schmid.";
 RL Am. J. Hum. Genet. 54:169-178(1994).
 RN [10]
 RP VARIANT SMCD ARG-591.
 RX MEDLINE=94272470; PubMed=8004099;
 RA McIntosh I., Abbott M.H., Worman M.L., Olsen B.R., Francomano C.A.;
 RT "Additional mutations of type X collagen confirm COL10A1 as the
 Schmid metaphyseal chondrodysplasia locus.";
 RL Hum. Mol. Genet. 3:303-307(1994).
 RN [11]
 RP VARIANT SMCD VAL-618.
 RX MEDLINE=95181449; PubMed=7876225;
 RA Chan D., Cole W.G., Rogers J.G., Bateman J.F.;
 RT "Type X collagen multimer assembly in vitro is prevented by a Gly618
 to Val mutation in the alpha 1(X) NCI domain resulting in Schmid
 metaphyseal chondrodysplasia.";
 RL J. Biol. Chem. 270:4558-4562(1995).
 RN [12]
 RP VARIANTS SMCD ARG-545; GLU-595; HIS-597; LYS-617; ARG-644 AND GLY-648.
 RX MEDLINE=95331767; PubMed=7607655;
 RA Bonaventure J., Chamande F., Maroteaux P.;
 RT "Mutations in three subdomains of the carboxy-terminal region of
 collagen type X account for most of the Schmid metaphyseal
 dysplasias.";
 RL Hum. Genet. 96:58-64(1995).
 RN [13]

RP VARIANT SMCD PRO-600.
 RX MEDLINE=96375754; PubMed=8782043;
 RA Wallis G.A., Rash B., Sykes B., Bonaventure J., Maroteaux P.,
 RA Zabel B., Wynne-Davies R., Grant M.E., Boot-Handford R.P.;
 RT "Mutations within the gene encoding the alpha 1 (X) chain of type X
 collagen (COL10A1) cause metaphyseal chondrodysplasia type Schmid but
 not several other forms of metaphyseal chondrodysplasia.";
 RL J. Med. Genet. 33:450-457(1996).
 RN [14]
 RP VARIANTS SMCD GLU-18 AND ARG-18.
 RX MEDLINE=97220591; PubMed=9067753;
 RA Ikegawa S., Nakamura K., Nagano A., Haga N., Nakamura Y.;
 RT "Mutations in the N-terminal globular domain of the type X collagen
 gene (COL10A1) in patients with Schmid metaphyseal
 chondrodysplasia.";
 RL Hum. Mutat. 9:131-135(1997).
 RN [15]
 RP VARIANTS SMD GLU-595.
 RX MEDLINE=99057503; PubMed=9837818;
 RA Ikegawa S., Nishimura G., Nagai T., Hasegawa T., Ohashi H.,
 RA Nakamura Y.;
 RT "Mutation of the type X collagen gene 'COL10A1' causes
 spondylometaphyseal dysplasia.";
 RL Am. J. Hum. Genet. 63:1659-1662(1998).
 RN [16]
 RP VARIANTS SMCD CYS-597.
 RX MEDLINE=99069781; PubMed=9852679;
 RA Sawai H., Ida A., Nakata Y., Koyama K.;
 RT "Novel missense mutation resulting in the substitution of tyrosine by
 cysteine at codon 597 of the type X collagen gene associated with
 Schmid metaphyseal chondrodysplasia.";
 RL J. Hum. Genet. 43:259-261(1998).
 CC -1- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC
 CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
 MINERALIZATION ZONES OF HYALINE CARTILAGE.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- PRIM: PROLINS AT THE THIRD POSITION OF THE TRIPETIDE REPEATING
 UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SCHMID TYPE
 METAPHYSEAL CHONDRODYSPLASIA (SMCD), A DOMINANTLY INHERITED
 DISORDER OF THE OSSEOUS SKELETON. THE CARDINAL FEATURES OF THE
 PHENOTYPE ARE MILD SHORT STATURE, COXA VARA AND A WADDLING GAIT.
 CC RADIOGRAPHY USUALLY SHOWS SCLEROSIS OF THE RIBS, FLARING OF THE
 METAPHYSES, AND A WIDE IRREGULAR GROWTH PLATE, ESPECIALLY OF THE
 KNEES.
 CC -1- DISEASE: DEFECTS IN COL10A1 ARE THE CAUSE OF SPONDYLOMETAPHYSEAL
 DYSPLASIA (SMD). SMD COMPRISES A HETEROGENEOUS GROUP OF HERITABLE
 SKELETAL DYSPLASIAS CHARACTERIZED BY MODIFICATIONS OF THE
 VERTEBRAL BODIES OF THE SPINE AND METAPHYSES OF THE TUBULAR BONES.
 CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
 CC -1- SIMILARITY: CONTAINS 1 C10 DOMAIN.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: X60382; CAA42933.1; -;
 DR EMBL: X63120; CAA46236.1; -;
 DR EMBL: X98568; CAA67178.1; -;
 DR EMBL: AL121963; CAB87590.1; -;
 DR EMBL: S68531; AAC60615.1; -;
 DR EMBL: X58879; CAA41686.1; -;
 DR EMBL: M74050; AAA61221.1; -;
 DR EMBL: X72579; CAA51170.1; -;
 DR EMBL: X72580; CAA51170.1; JOINED.
 DR PIR: S15826; S15826.
 DR PIR: S30086; S30086.
 DR PIR: A43901; A43901.
 DR PIR: S18249; S18249.

DR PIR; S21856; S21856.
 DR PIR; S26396; S26396.
 DR MIM; 120110; -.
 DR MIM; 165500; -.
 DR MIM; 184250; -.
 DR InterPro; IPR001073; Clq.
 DR InterPro; IPR000087; Collagen.
 DR Pfam; PF00386; Clq; 1.
 DR Pfam; PF01391; Collagen; 7.
 DR PRINTS; PRO0007; Collagen; 1.
 DR SMART; SM00110; Clq; 1.
 DR PROSITE; PS01113; Clq; 1.
 DR Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 DR Cartilage; Collagen; Signal; Disease mutation; Polymorphism.
 FT SIGNAL 1 18
 FT CHAIN 19 680 COLLAGEN ALPHA 1(X) CHAIN.
 FT DOMAIN 19 56 NONHELICAL REGION (NC2).
 FT DOMAIN 57 519 TRIPLE-HELICAL REGION.
 FT DOMAIN 520 680 NONHELICAL REGION (NC1).
 FT DOMAIN 545 680 Clq.
 FT VARIANT 18 18 G-> E (IN SMCD).
 FT VARIANT 18 18 /FTID-VAR_001838.
 FT VARIANT 18 18 G-> R (IN SMCD).
 FT VARIANT 545 545 /FTID-VAR_001839.
 FT VARIANT 591 591 /FTID-VAR_001840.
 FT VARIANT 591 591 C-> R (IN SMCD).
 FT VARIANT 595 595 /FTID-VAR_001841.
 FT VARIANT 595 595 G-> E (IN SMCD AND SMD).
 Query Match 3.9%; Score 110.5; DB 1; Length 680;
 Best Local Similarity 30.7%; Pred. No. 1.2; Indels 35; Gaps 7;
 Matches 43; Conservative 11; Mismatches 51; Indels 35; Gaps 7;
 QY 412 PYRGDTG--GASSEDSTLSTSLPGP--KPGAFPGNHVAGSGSLPPPALCGASACD 467
 DB 463 PSKGDPGSPGPPGACITKGLNGTGPGRPGHSGEGC---LPPPP----- 510
 QY 468 VSVRVVGEPTPEARVPGRGICIDLAIDLSAPLLSOVAPSLFNGSIVQLSQSVTAYVSA 527
 DB 511 -----GPPGP---PGQ-----AVMEGEFIKAGORPSLSCTPLVSANOGVTGMPVSA 553
 QY 528 AGLGVAIR--FATGVPEK 545
 DB 554 FTVILSKAIPATGPIPFK 573
 RESULT 14
 GALT_HUMAN
 ID GALT_HUMAN STANDARD; PRT; 368 AA.
 AC 060755;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Galanin receptor type 3 (GAL3-R) (GALR3).
 GN GALR3 OR GALNR3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eularchia; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9838766; PubMed=9722565;
 RA Smith K.E., Walker M.W., Artymyshyn R., Bard J., Borowsky B.,
 RA Tamm J.A., Yao W.-J., Vayssie P.J.-J., Branchek T.A., Gerald C.,
 RA Jones K.A.;
 RT "Cloned human and rat galanin GALR3 receptors: pharmacology and
 RT activation of G-protein inwardly rectifying K+ channels.";
 RL J. Biol. Chem. 273:23321-23326(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 CC Bennett M.M., Lescoe M.K., Gallipoli P.Z., Ramabhadran T.V.;

RT "Homologue of the human galanin 2 receptor gene isolated from a human
 RT uterus cDNA library.";
 RT submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99048961; PubMed=9832121;
 RA Kolakowski L.F. Jr., O'Neill G.P., Howard A.D., Broussard S.R.,
 RA Sullivan K.A., Feighner S.D., Sawdzargo M., Nguyen T., Kargman S.,
 RA Shiao L.-L., Hreniuk D.L., Tan C.P., Evans J., Abramowitz M.,
 RA Chateaufort A., Coulombe N., Ng G., Johnson M.P., Thattian A.,
 RA Khasabov H., George S.R., Smith R.G., O'Dowd B.F.;
 RT "Molecular characterization and expression of cloned human galanin
 RT receptors GALR2 and GALR3.";
 RL J. Neurochem. 71:2239-2251(1998).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9912120; PubMed=9928159;
 RA Iismaa T.P., Fathi Z., Hort Y.J., Iben L.G., Dutton J.L., Baker E.,
 RA Sutherland G.R., Shine J.;
 RT "Structural organization and chromosomal localization of three human
 RT galanin receptor genes.";
 RL Ann. N.Y. Acad. Sci. 863:56-63(1998).
 RN [5]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Baare D.M.,
 RA Clamp M., Smink L.J., Alnscough R., Almeida J.P., Babbage A.,
 RA Baguley C., Bailey J., Barlow K., Bates K.N., Beasley O., Bird C.P.,
 RA Blakey S., Bridgeman A.M., Buck D., Burgess J., Burrill W.D.,
 RA Burton J., Carder C., Carter N.P., Chen Y., Clark G., Clegg S.M.,
 RA Cobley V., Cole C.G., Collier R.E., Connor R.E., Conroy D., Corby N.,
 RA Coville G.J., Cox A.V., Davis J., Dawson E., Dhani P.D., Dockree C.,
 RA Dodsworth S.J., Durbin R.M., Ellington A., Evans K.L., Fey J.M.,
 RA Fleming K., French L., Garner A.A., Gilbert J.G.R., Goward M.E.,
 RA Grahame D., Griffiths M.N., Hall C., Hall R., Hall-Tamlyn G.,
 RA Heathcote R.W., Ho S., Holmes S., Hunt S.E., Jones M.C., Kershaw J.,
 RA Kimberley A., King A., Laird G.K., Langford C.F., Leversha M.A.,
 RA Lloyd C., Lloyd D.M., Martyn I.D., Mashreghi-Mohammadi M.,
 RA Matthews L., Mccann O.T., Mcclellan J., McLaren S., McMurtry A.A.,
 RA Milne S.A., Mortimore B.J., Odell C.N., Pavitt R., Pearce A.V.,
 RA Pearson D., Phillimore B.J., Phillips S.H., Plumb R.W., Ramsay H.,
 RA Ramsey Y., Rogers L., Ross M.T., Scott C.E., Sehra H.K., Skuce C.D.,
 RA Smalley S., Smith M.L., Soderlund C., Spragon L., Steward C.A.,
 RA Sulston J.E., Swann R.M., Vautin M., Wall M., Wallis J.M.,
 RA Whiteley M.N., Willey D., Williams L., Williams S., Williamson H.,
 RA Wilmer T.E., Wilming L., Wright C.L., Hubbard T., Bentley D.R.,
 RA Beck S., Rogers J., Shimizu N., Minoshima S., Kawasaki K., Sasaki T.,
 RA Asakawa S., Kudoh J., Shintani A., Shibuya K., Yoshizaki Y., Aoki N.,
 RA Mitsuyma S., Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S.,
 RA Do A., Do T., Dorman A., Fang F., Fu Y., Hua A., Kenton S.,
 RA Lai H., Lao H.I., Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E.,
 RA Nguyen T., Pan H., Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S.,
 RA Sloan D., Song L., Wang Q., Wang Y., Wang Z., White J., Williamson D.,
 RA Wu H., Yao Z., Zhan M., Zhang G., Chisoso M., Murray J., Miller N.,
 RA Mux P., Fulton R., Johnson D., Bemis G., Bentley D., Birdshaw H.,
 RA Bourne S., Cordes M., Du Z., Fulton L., Goela D., Graves T.,
 RA Hawkins J., Hinds K., Kemp K., Latreille P., Layman D., Ozersky P.,
 RA Rohlfing T., Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K.,
 RA Nelson J., Korf I., Bedell J.A., Hillier J.A., Merdes E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saita S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edelmann L., Kim U.J., Shizuya H., Simon M.I., Dunanski J.P.,
 RA Peyrard M., Kedra D., Seroussi E., Fransson I., Tapia I., Brider C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenleitch A., Hartman K., Hir X.,
 RA Khan A.S., Lane L., Tilihan Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 CC -!- FUNCTION: RECEPTOR FOR THE HORMONE GALANIN.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC
DR EMBL; AF073739; AAC35944.1; -
DR EMBL; AF067733; AAC18860.1; -
DR EMBL; AF129514; AAD47348.2; -
DR EMBL; AF129513; AAD47348.2; JOINED.
DR EMBL; 297630; CAB42831.1; -
DR MIM; 603692; -
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_HODOPSIN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS50262; G-PROTEIN_RECEP_FL_2; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein;
KM Multigene family; 14poprotein; Palmitate.
FT DOMAIN 1 20 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 1 41 1 (POTENTIAL).
FT DOMAIN 42 57 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 58 78 2 (POTENTIAL).
FT DOMAIN 79 96 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 97 118 3 (POTENTIAL).
FT DOMAIN 119 138 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 139 159 4 (POTENTIAL).
FT DOMAIN 160 184 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 185 205 5 (POTENTIAL).
FT DOMAIN 206 236 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 237 257 6 (POTENTIAL).
FT DOMAIN 258 259 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 260 280 7 (POTENTIAL).
FT DOMAIN 281 368 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 6 6 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 95 172 BY SIMILARITY.
FT LIPID 308 308 PALMITATE (BY SIMILARITY).
FT DOMAIN 218 225 POLY-ALA.
FT DOMAIN 310 318 POLY-ARG.
SQ SEQUENCE 368 AA; 39573 MW; E8F46B642C3150 CRC64;

Query Match 3.8%; Score 110; DB 1; Length 368;
Best Local Similarity 21.9%; Pred. No. 0.71;
Matches 94; Conservative 35; Mismatches 145; Indels 156; Gaps 17;

QY 87 PIVALSIGILSLFLPRAG--WLAGLCPDPRPLE-----LALLIGVGLDFCGQY 138
DB 21 PIVFAL-----IFLGTGNGVLAVLLOPGPSAKOBPGSTTDLTINTLAVDLCFTL 73
QY 139 CFTPLEALLSDFRDPDHCRAVSVAVMISIGGICGYLLPAIDWDTSLAPYLCTOECC 198
DB 74 CCVPRQAVT-----YTLDN--W----- 88
QY 199 LFGLLT-----LIFLTCVAATLLVAEBAALGPEPABGLSAPSLSPHCCPCRRARLARFN 252
DB 89 LEGALVCAVHLLVLTMYASSFT-----LAASVDRY-----LAAY- 125
QY 253 LGALLPRLHOLCCMRPRLRLRFVALLCWMALMTFTLFYDFVEGGLYOGVRAPEPGE 312
DB 126 -----HPLRSRLKTRPRNARAAGVGLWLLALFSAPIYSLTYGVRYGALCIVPAME 177
QY 313 ARRAYDEGVKMSGLFLQCA-ISLVFSIVMDRLVQRF-----TRAVYLASYAA 361
DB 178 DARRRALDVATPAAGYLLPVANVSLAYGRTLRFMAAAGPAGAAAARARRATGRAGRAM 237
QY 362 PPVAA-----GATCLSHSVAVVTASNAALTGFFSALQILPYTLASLY 403
DB 238 LAVALVALCWMGPHHALILCEWYGRFAFSPATYACRLASHCLAYANSCINPLVYALASHR 297
QY 404 -----HREKQVFLPKYRGDTCGASSEDLSMISFLDPRGAPFPNGHVH 447
DB 298 FRARPRRLMPCGRRRRRRARRA-LRRVPRASSG-----PPGCGPDARPSGRLL 344

QY 448 AGGSGILPPP 457
DB 345 AGG-CQGP 353

RESULT 15
NOOC_THETH
ID NOOC_THETH STANDARD; PRT; 606 AA.
AC 056227;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE NADH-ubiquinone oxidoreductase chain 12 (EC 1.6.5.3) (NADH
DE dehydrogenase 1, chain 12) (NDH-1, chain 12).
GN NOO12.
OS Thermus aquaticus (subsp. thermophilus)
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.
OX NCBI_TaxID=274;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=HB8 / ATCC 27634;
RX MEDLINE=97172490; PubMed=9020134;
RA Yano T., Chu S.S., Sled V.D., Ohnishi T., Yagi T.;
RT "The proton-translocating NADH-quinone oxidoreductase (NDH-1) of
RT thermophilic bacterium Thermus thermophilus HB-8. Complete DNA
RT sequence of the gene cluster and thermostable properties of the
RT expressed NOO2 subunit.";
RL J. Biol. Chem. 272:4201-4211(1997).
CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.
CC -1- SUBUNIT: COMPOSED OF 14 DIFFERENT SUBUNITS. SUBUNITS NOO7-14
CC -1- SUBUNIT: THE MEMBRANE SECTOR OF THE COMPLEX.
CC -1- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE
CC OF CHLOROPLASTS OR MITOCHONDRIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

Query Match 3.8%; Score 110; DB 1; Length 606;
Best Local Similarity 21.4%; Pred. No. 1.2;

| | | | |
|----|-----|---|-----|
| QY | 62 | VLGIWVAVLLGASADHNRGRGRG--RRPFMAISLIGLISLEFLP-----RAGN | 108 |
| Db | 3 | LLGLTLPLPLGFA---LLGLGFKRMREPLGCVLASGLVASFLGAGLLSGGARROAEN | 59 |
| QY | 109 | LAGL---LCDDPRLLEALLLIGVGLLDFFCGQVCFTPEALLSDLEFDPDHCROAYSU- | 164 |
| Db | 60 | LPGLPFSLLDNLGSMFLIYLVGFLIHVAIYG-----MGDDPGYSR--FPAYF | 108 |
| QY | 165 | ---AFMISLGGCGLLPLPAIDWDTSLAFLYLGIOECLFG----- | 201 |
| Db | 109 | NLFAMITLVLADSYRMFGWGVGLASF-----LLGFMYKNPOYADSARKAFTVN | 162 |
| QY | 202 | -----LLTLVFLTCVAATLVVE--EALAGPTEPAEGISAPLSLPHCCRCARLAFRN | 252 |
| Db | 163 | RIGDGFMYGMAIIMALYGTISISELKEAMGR-----LKNPL-----LALAG | 206 |
| QY | 253 | LGALLPRLHQCSCMRPTLRRLFVAELCSWMAIMTFLPYDFVGEQ-----LYQGVRAE | 308 |
| Db | 207 | L-----LRLGANGKSAQIRIMWTLRDM | 230 |
| QY | 309 | PGTARRKHIDEGVMSGSLGFLQCAISLVSLWDLRYQRRGTRAVULASVAARVAA-G | 367 |
| Db | 231 | AGPFPVSALIHAAFMVAGVULLIRSSFSLVLPD---VSAIIVAGLLTAAAGALSFG | 287 |
| QY | 368 | ATCISHVAVVUTASAAALGFESALQILPRTLASLAREKQV----- | 410 |
| Db | 288 | QTDIKKIYASTISO--LGVETFLAAGGAYWA-LFHVFTFAEFKALLFTLASGSVINALG | 344 |
| QY | 411 | -----LPKYRQDT-----CGASSEDLSMTSFLPGPKGAPRN | 443 |
| Db | 345 | GEQDVYRKMGJMKHLRQYRMHALLGALALGSLPLDSFWSKDAITLTLVPPGGVGF-- | 402 |
| QY | 444 | GHVAGGSGGLLPPRPALCGASACVSVRVVVG-----PTEARV---PGRGICLDLAI | 494 |
| Db | 403 | -YVGA-----LL--VAVLTAAMYAMFWFLVLEGERGHNHREARVMIWENHNLIALG-SV | 454 |
| QY | 495 | LDSAPILSQVAPSLFMGSIYVLSQSVTAUWV--AAGGIVAIYFPAQVU | 542 |
| Db | 455 | LAGTIALPHRLPRLVERPLKALALEVENHNISLGAEMGILTALSAAVALL | 503 |

Search completed: June 26, 2002, 13:48:18
Job time: 401 sec

5 PAGE BLANK (USP 12)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 12:40:07 : Search time 70.32 Seconds
(without alignments)
755.651 Million cell updates/sec

Title: US-09-838-785-2

Perfect score: 2861

Sequence: 1 MVQRLWVSRLRRKQAQLL.....AIYFATQVFDKSLAKYSA 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------------|
| 1 | 347.5 | 12.1 | 515 | 2 | T14340 sucrose-proton tra |
| 2 | 337.5 | 11.8 | 523 | 2 | T12198 sucrose transport |
| 3 | 330 | 11.5 | 516 | 2 | JQ2389 sucrose transport |
| 4 | 328.5 | 11.5 | 594 | 2 | G84441 probable sucrose-p |
| 5 | 324.5 | 11.3 | 525 | 2 | S28052 sucrose transport |
| 6 | 323.5 | 11.3 | 512 | 2 | S38196 sucrose transport |
| 7 | 323.5 | 11.3 | 512 | 2 | G86360 probable sucrose-p |
| 8 | 317.5 | 11.1 | 507 | 2 | S48789 sucrose transport |
| 9 | 311 | 10.9 | 533 | 2 | S43142 sucrose transport |
| 10 | 305.5 | 10.7 | 531 | 2 | T14339 sucrose-proton tra |
| 11 | 297.5 | 10.4 | 428 | 2 | S48788 sucrose transport |
| 12 | 294 | 10.3 | 510 | 2 | S38657 sucrose transport |
| 13 | 294 | 10.3 | 523 | 2 | S51114 sucrose-proton sym |
| 14 | 292.5 | 10.2 | 512 | 2 | E86741 probable sucrose t |
| 15 | 292 | 10.2 | 537 | 2 | T02982 probable sucrose t |
| 16 | 286 | 10.0 | 474 | 2 | A86234 hypothetical prote |
| 17 | 281.5 | 9.8 | 491 | 2 | E86691 probable sucrose-p |
| 18 | 279.5 | 9.8 | 513 | 2 | S38197 sucrose transport |
| 19 | 276.5 | 9.7 | 492 | 2 | A84520 probable sucrose-p |
| 20 | 273 | 9.5 | 503 | 2 | S52377 sucrose transport |
| 21 | 235.5 | 8.2 | 553 | 2 | T38541 probable sucrose c |
| 22 | 191 | 6.7 | 452 | 2 | F75217 hypothetical prote |
| 23 | 187.5 | 6.6 | 541 | 2 | B87532 transporter, proba |
| 24 | 142 | 5.0 | 544 | 2 | S75696 melibiose carrier |
| 25 | 139.5 | 4.9 | 454 | 2 | A75444 hypothetical prote |
| 26 | 138.5 | 4.8 | 430 | 2 | E75217 transporter PAB217 |
| 27 | 136 | 4.8 | 418 | 2 | B87536 membrane protein, |
| 28 | 133 | 4.6 | 422 | 2 | G83413 probable MFS trans |
| 29 | 128.5 | 4.5 | 422 | 2 | G83503 probable MFS trans |

| | | | | | |
|----|-------|-----|-----|---|----------------------------|
| 30 | 128.5 | 4.5 | 594 | 2 | A83096 probable permease |
| 31 | 127.5 | 4.5 | 472 | 2 | F82639 resistance protein |
| 32 | 127 | 4.4 | 451 | 2 | E81781 probable integral |
| 33 | 126 | 4.4 | 394 | 2 | G98213 drug efflux protei |
| 34 | 126 | 4.4 | 394 | 2 | A83073 tetracycline resis |
| 35 | 126 | 4.4 | 451 | 2 | A81206 sugar transporter, |
| 36 | 125.5 | 4.4 | 399 | 2 | JQ1479 tetracycline resis |
| 37 | 125 | 4.4 | 391 | 2 | S74688 hypothetical prote |
| 38 | 125 | 4.4 | 397 | 2 | H75515 tetracycline-efflu |
| 39 | 124.5 | 4.4 | 309 | 1 | S38656 tetra protein - pse |
| 40 | 123.5 | 4.3 | 400 | 2 | D97643 probable efflux pr |
| 41 | 123.5 | 4.3 | 400 | 2 | AG2866 MFS permease [drug |
| 42 | 123 | 4.3 | 440 | 2 | AD0986 hypothetical metab |
| 43 | 121.5 | 4.2 | 398 | 2 | C90349 multidrug-efflux t |
| 44 | 121.5 | 4.2 | 503 | 2 | S43017 puromycin resistan |
| 45 | 121 | 4.2 | 483 | 2 | AF2363 hypothetical prote |

ALIGNMENTS

| Query Match | 12.1% | Score 347.5 | DB 2 | Length 515 |
|-----------------------|---|-----------------|---|------------|
| Best Local Similarity | 25.9% | Pred. No. 1e-18 | | |
| Matches 131 | Conservative 80 | Mismatches 215 | Indels 79 | Gaps 17 |
| QY 17 | QLLLVNLTFGLVEVCLAGITVPPLLLEGVGEKFMVNLGIGVPLGVPLGASD 76 | DB 179 | PAIDWTSALAPYLGTQEECLFGLTLTFLTCVAATLLVAEEALGPTPAEGLSPSIS 238 | |
| DB 34 | KLVVAIAIAGVQFGWALQSLTLPVYQLGIPHKMAAVIWLGPISGMLVQPIVGYSD 93 | DB 211 | PS--KTHACDLYCANLKSCFIISALLIITIVVALSVRENS--GPPDADAEP--- 263 | |
| QY 77 | HMRGRGRRRPPTWALSILSLIFLIPRAGWLAGL-----LCPPRPPLALILIGV 129 | DB 94 | HCSSGRRRRPPTASGAGCAISVILI---GPAADISYRAGDMSKTLPRATVVEIGF 150 | |
| QY 130 | GILDFGGOVCPPLLEALLSDLFR-DPDHCRQAVSYVAFMISLGGCGY-----LL 178 | DB 151 | WILDVANNMLQGPCRALDLDCSGDTRMRMSANAFSFMVAGNIIIGVAGSYNNLYKLF 210 | |
| QY 179 | PAIDWTSALAPYLGTQEECLFGLTLTFLTCVAATLLVAEEALGPTPAEGLSPSIS 238 | DB 211 | PS--KTHACDLYCANLKSCFIISALLIITIVVALSVRENS--GPPDADAEP--- 263 | |
| QY 239 | PHCCPCRARLARLNRIGALLPRHLQCCMRPTRLRLFAVALCSWMAIMFTLFYDFVGE 298 | DB 264 | ---PSSGRTPV--FGELIGALKDL---PRPWLILIVTCLNMWIMPFILFDTPWMSR 313 | |
| QY 299 | GLYGVRAPEPTEARHRYDEGVMSLGFLOCAISLVEFSLVMDRLVORFGRVAYYLS 358 | DB 314 | EYGGT--AGCG---KLYDQVRAGALGLLLNSVYLGITSLAVEYLVAGGVKLI-LWG 366 | |
| QY 359 | VAFVPAAGATCU-----SHSV-----AVVTASA---ALTGFFSALQTL 395 | | | |

Db 367 FVNFILALIGLVMTVVVSKVAQHOREHSANGOLLPPSAGVAGALSLFSLICPISTITYSI 426

QY 396 PPTLASLYHREKOVFLPKRYGDTGASSEDLSMTSPFGPKPG-----APPNGHYGA 448

Db 427 PPLASTSYSSGAGGSLSLGVLNLAIVPQMIVSVLAFGFDLSLGGGNLPAFVYGAISA 486

QY 449 GSGGLP-----PPALCGASACDVS 469

Db 487 AISGVLAIVLLPKPSKDAASKLSIS 511

RESULT 2

T12198
sucrose transport protein - fava bean
C:Species: Vicia faba (fava bean)
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 20-Jun-2000
C:Accession: T12198
R:Weber, H.; Borisjuk, L.; Helm, U.; Sauer, N.; Wobus, U.
Plant Cell 9, 895-908, 1997
A:Title: A role for sugar transporters during seed development: molecular characterization
A:Reference number: 217451; PMID:97355984
A:Accession: T12198
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-523 <WEB>
A:Cross-references: EMBL:293774
C:Genetics:
A:Gene: sut
C:Superfamily: common tobacco sucrose transport protein

Query Match 11.8%; Score 337.5; DB 2; Length 523;
Best Local Similarity 25.8%; Pred. No. 6e-18;
Matches 134; Conservative 76; Mismatches 185; Indels 125; Gaps 20;

QY 17 QLLVNLITFLEVCIAAGITVYPPLLEVEEKEFMVIGIPVGLVGPVLLGSASD 76

Db 36 KIMVVASIAAGVQFGMALQSLTPPYVOLLGIHHTMAVYIWLCPISGMLVQPIVGYHSD 95

QY 77 HMRGRVRRPFIWALSGILSLFLIPRAGMLAGL-----LCPPRPLEALLILGV 129

Db 96 RCTSRFRGRFFIAAGSTAVAINVFLI---GYADLGHSPDSLDQKRPRAIGIFVYVF 152

QY 130 GLDDECGVCFPLEALLSLDFR-DPDHCRQAVSYVAFMISLGGCLGYLLPAID----- 182

Db 153 WILDVANMMLGCPRALIGDLGACNQKRTNANAFSEFMAVGNVLYGAAGAYSKLYHF 212

QY 183 --WQTSALAPYLQOEELFGLTLIFLTQVAATLLV-AEEAALGPTEPA-----EGLSAP 235

Db 213 PFTKACNVYCANLKSCF--LSTALTTLVATSLIYKETAALPEKTVVTTEDGSSG 270

QY 236 SLSPHCCPCRRARL--AFRNILGALLPRLHQLCCRMPTLRLLFVAELCSMALMPTFLPYT 293

Db 271 GM-----PCFQOLSGAFREL-----KRPWMLLLVCLMWIMAFPLLPDPT 311

QY 294 DFVEGELGYGVPAEPGETEARRH-YDEGVKMGSLGLFLOCAISLVSLVMDRLVQFGR 352

Db 312 DMNGKEVY-----GGTVGEGHAIDMGVRGALGLMLNSVVLGATSLGVDIRAGV-G 363

QY 353 AVYLASVAAPFAAGATCISHVAVY-----TASALTGF 387

Db 364 VKRLMGIVNFL--AICLGLTLVTKLAQHSROYAPGTGALGDLPPSEGIKAGALLTF 420

QY 388 TFSALQI-----LPYTLASLYHREKOVFLPKRYGDTGASSEDLS-----MTS 430

Db 421 SVLGVPPLAITYSIPPLASISF-----SSTSGAGGSLSLGVNLATIVPQMFVS 468

QY 431 FLPGPKPG-----APPNGHYGAGSG-----LLPPP 458

Db 469 VLSGPMDALFGGGLNLPFAFVYGAVALASGIISLILPSP 508

RESULT 3

JQ2389
sucrose transport protein - potato
C:Species: Solanum tuberosum (potato)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 21-Jul-2000
C:Accession: JQ2389; S40310
R:Riesmeier, J.W.; Hirner, B.; Frommer, W.B.
Plant Cell 5, 1591-1598, 1993
A:Title: Potato sucrose transporter expression in minor veins indicates a role in phl
A:Reference number: JQ2389; PMID:94146554
A:Accession: JQ2389
A:Molecule type: mRNA
A:Residues: 1-516 <RIE>
A:Cross-references: EMBL:X69165; NID:9439293; PIDN:CAA48915.1; PID:9439294

A:Experimental source: cv. Desiree
C:Comment: The gene encoding for this protein is highly expressed in mature leaves.
C:Superfamily: common tobacco sucrose transport protein
C:Keywords: glycoprotein; transmembrane protein
F:31-53/Domain: transmembrane #status predicted <TM1>
F:67-86/Domain: transmembrane #status predicted <TM2>
F:103-122/Domain: transmembrane #status predicted <TM3>
F:141-160/Domain: transmembrane #status predicted <TM4>
F:180-200/Domain: transmembrane #status predicted <TM5>
F:226-248/Domain: transmembrane #status predicted <TM6>
F:285-304/Domain: transmembrane #status predicted <TM7>
F:331-349/Domain: transmembrane #status predicted <TM8>
F:366-385/Domain: transmembrane #status predicted <TM9>
F:409-427/Domain: transmembrane #status predicted <TM10>
F:429-448/Domain: transmembrane #status predicted <TM11>
F:3.92/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 11.5%; Score 330; DB 2; Length 516;
Best Local Similarity 24.7%; Pred. No. 2.2e-17;
Matches 132; Conservative 93; Mismatches 197; Indels 112; Gaps 21;

QY 4 RLWYSRLLRHRRKAOLLVNLTFFGLEVCIAAGITVYPPLLEVEEKEFMVIGIPVL 63

Db 29 KLV-----KTIYVASIAGVQFGMALQSLTPPYVOLLGIHHTMAVYIWLCPISGMLVQPIVGYHSD 78

QY 64 GLVCPVLLGSASDHMRGRYRRRPFIWALSGILSLFLIPRAGMLAGL-----LCPP 116

Db 79 GMIVQPVVGYYSNDKSSFRGRRPFIAGALVMAIVFLI---FFAADLGHASGDTLEKG 135

QY 117 PRPLEALLIIGVGLDPCGVCFPLEALLSLDFRQD-HCRQAVSYVAFMISLGGCLG 175

Db 136 FKPRAIAYVVFVWILDVANNNMLGCPRALIADISGKSGMRRTANAFSEFMAVGNILG 195

QY 176 YLLPAIDW-----DTSALAPYLQOEELF-GILTLIFLTQVAATLLVAEE-AALGP 225

Db 196 YVAGSYSHLFYFPPSKKACDMVCANLKSCFFIAPILLSITTLATLVRENELPERKE 255

QY 226 TEPAGLSAPSLSPHCCPCRRARLAFRNILGALLPRLHQLCCRMPTLRLLFVAELCSMAL 285

Db 266 QEIDELKLAGAG-----KSKVPF--FGEIFGALKEI-----PRPMWMLLVTCGLMTAW 301

QY 286 MTFITFYTDYFGEGLGYGVPAEPGETEARRH-YDEGVKMGSLGLFLOCAISLVSLVMDRL 345

Db 302 PFFFLYTDYDMAKKEVGG--QVGD--ARLYDLSVRGAGKGLLQSYVLYGPMISLGVFL 355

QY 346 VQFEGTRAVYLASVAAPFAAGATCISHVAVY-----TASALTGF 388

Db 356 GKRIQ-GAKRLMGLINFL--AICLMTLITVTMAEKSRQHDPAITMGPTPGVIGAL 411

QY 389 --FSAIQI-----LPYTLASLYHREK-----QVFLPKRYGDTGASSEDLS 426

Db 412 LIFALGIPPLAATFISIPPLASISFSSNRSGGSLGVNLATIVPQMLVSLVYGGPMWDL 471

QY 427 LMTSFLPGPKGAPFPNGHYGAGSG-----LLPPPALCGASACDVSVRYVY 475

Db 472 FCGGNLPG-----FVVGAVAAASAVALATLTLPLSPPA-----DAKPAVAMG 512

RESULT 4

| | | | | |
|---------------------------|--------|--------------------|-------------|-------------|
| Query March | 11.5% | Score 328.5; | DB 2, | Length 594; |
| Best Local Similarity | 24.1%; | Pred. No. 3.3e-17; | | |
| Matches 120; Conservative | 72; | Mismatches 174; | Indels 131; | Gaps 18 |

RESULT 5
S28052

A:Reference number: S28052; MUID:93099843
A:Molecule type: mRNA
A:Residues: 1-525 <RIB>
A:Cross-references: EMBL:X67125; NID:g21318; PIDB:CAA47604.1; PID:g21319
C:Superfamily: common tobacco sucrose transport protein
C:Keywords: transmembrane protein

Query Match 11.3%; Score 324.5; Db 2; Length 525;
Best Local Similarity 23.2%; Pred.No.5,8e-17;
Matches 129; Conservative 91; Mismatches 196; Indels 141; Gaps 20;

OY 23 LITFGLNCLAGITY-----VPLLLEVEEKEFTMWLGIPVLGCVPLLSGA 74
| | :
Db 35 LKTLGISVASVAGVOFGMALQLSLTPVOLIGIPIPTMAATWLCGPISGMIVPPLGVY 94
OY 75 SDHMGGRARRPRFWALSLGILSFLIPRAGWL -GLLPDP----RPLEALLLT 127
| | :
Db 95 SRCTSRFGRRRPFPAAGALVAVALGT--GFADIGAASGDPPGNVAKPRALAIFEV 151
OY 128 GVGLDFCGOVCFPTLEALLSDLFR-DPHCRQASVYAFAIMSLGCGLGYLLPAID--- 182
| | :
Db 152 GWIIIDVANNTIOGCCRALLDAMAGSQTKRYANAFFSFNALINIGSYAAGYSRLYT 211
OY 183 ---WDTSALAPLYQTQEBCIEFLT-LLETCAATLIVAEALGPTEAEGLSABSL 237
| | :
Db 212 VEPFTKTAACDYYCANLKSCFEISTILLIVTILALSVYKERQITIDEIQEEEDLKRN 271
OY 238 SPHCPCRARLRFRNLGALLPRLHOLCCMRPTLRPLFAELCSMMALMTFTLFETDFVG 297
| | :
Db 272 SSGC---ARLPF--FGDLIALKLD---PKMLILLVTALNMWAMPPLLEFDTDWG 321
OY 298 ESLYGVRAPEPGTEARRHYDEGVBMGSGLFLCALISLVESLVMDRLVQRFG--TRAVY 355
| | :
Db 322 KEVYGGT-----VEBGKIYDGVIHAGALGMINSVLVGMSLSTEGLARVMYGAKKRIMG 375
OY 356 LASVAAPVAAGATCLSHSAVVYASALTGFTFSALQILPYTLASLYHREKVELPKYR 415
| | :
Db 376 IYNILL-----AVCLAMTV-LVTKSAE-----HFRDSHHIM---- 405
OY 416 GDTGGAASSDSLMTSFLPDKRKPARPNNGHVAGSGLLPPPALOGASACVSRYRVY 475
| | :
Db 406 -----SSAPPBPAA-GVKGALAIPLAVLG 429
OY 476 EPTEARV-VP-----GRGICDLAIDSASFILSQV-----AP 506
| | :
Db 430 IPLATFTSIPLAFASISSSGGGLSIGVNLIAIVPMVSVTSGPWDMFGGGINLP 489
OY 507 SLFMGSIQLQSQTAY 523
| | :
Db 490 AFGVGAVALATASAVLSLF 506

RESULT 6
S38196
sucrose transport protein SUC2 - Arabidopsis thaliana
N:Alternate names: sucrose-proton symporter SUC2
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Dec-2000
C:Accession: S38196; T00773
R:Sauer, N.K.
submitted to the EMBL Data Library, October 1993
A:Description: SUC1 and SUC2, two sucrose transporters from Arabidopsis thaliana.
A:Accession: S38196
A:Accession: S38196
A:Molecule type: mRNA
A:Residues: 1-512 <SAU>
A:Cross-references: EMBL:X75382; NID:g407091; PID:g407092
R:Vyotsakaia, V.S.; Schwartz, J.R.; Tortumii, M.; Yu, G.; Qi, O.; Kuan, A.; Liu, S.;
etz, D.; Li, Y.; Palm, C.J.; Shinn, P.; Sun, H.; Davis, R.W.; Eckert, J.R.; Federspiel,
Submitted to the EMBL Data Library, July 1998
A:Description: Arabidopsis thaliana chromosome 1 BAC T22J18 sequence, complete sequen

A:Reference number: Z14202
A:Accession: T00773
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-450, 'A', 492-512 <VS>
A:Cross-references: EMBL:AC003979; NID:g3172156; PID:g3287687; GSPDB:GN00059; ATSP:T22J1
A:Experimental source: cultivar Columbia
C:Genetics:
A:Gene: SUC2; ATSP:T22J18..12
A:Map position: 1
A:Introns: 419/3; 441/1; 455/3
C:Superfamily: common tobacco sucrose transport protein
C:Keywords: sugar transport

tions: 419/3.4441/1; 455/3
 preferably: common tobacco sucrose transport protein
 s: words: sugar transport
 ery Match 11.3%; Score 323.5; DB 2; length 512;
 st Local Similarity 24.1%; Pred. No. 6.7e-17;
 tches 123; Conservative 92; Mismatches 195; Indels 101; Gaps 18;

```

OY 14 RKAOLILVNLTLEFGEVLAAIGITVYVPLLEVEGEEMFMVIGISVGLIYVCPBLG 73
   1 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 28 RLKRIISVSIAGVQFPMALQLSILTPYVOLLPHPKMASLIMICGSIISGLVQPIYG 87
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 74 ASDHMRGRTYRRRPPIWALSIGILISFLIPRAGLACL-----LCPDRPELALIL 126
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 88 HSDRCTSRFRGRPRPIYVAGILYVAVELI---GYADIGHSMGDOLKPKRTAIAIFA 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 127 LGVGLDRCGVGCFFPLALLSDLFR-DPDHROAYSYAFAMISIGCGCY-----LL 178
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 145 LGFWLLDVANNITLQCPCAFPLADLSAGNAKKTRTANAFSPFMANGVNLVGAAGSYRNLX 204
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 179 PAIDMD-TSALAPYLCTOECLFGLLTLLIPLTCAATLLVVEEALGPTPEAGLSBSL 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 205 KYVPTMTESCDLYCANAKCTCFISITLL-LIVTFVSLCYEKEKPMTEPEPTADKCA--- 259
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 238 SPHCPCBARLAFNRILGALLPRLHOLCCMRPTLRLFYAELCSVMALMTFLFTDPCV 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 260 -----SNVFF--FGELFGAFKEL---KRPMLLILYALNLIAMFPFLFTDMDWG 305
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 298 EGLYGVRAEPETEARRHRYDEGVBMGSLGLELQACALISVLSLYMDRLVQRFGRAYVLA 357
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 306 REVYGGNSDAPRTAASKKILYNDGVABGALGLMLNLIYLGFMISIGWIGRKILG-CAKRLM 364
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 358 SYAAPVPAAGALCLSHSAV-----VTASA-----ALTG-----FT 388
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 365 GIVNFIIL--ALCIAMTYVYVTKQAEHRRHDGCAKTPPGNVTAGALLTFALLGIPQAIT 421
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 389 FSALQILLEYTIALSLHREK-----QVFLPYRGDTGGASSDSIMTSFLDPGP 435
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 422 FS-----IPFALASIFSTNSGAGQGLSLGVLNLIAYVPPQWIVISVGGPFDLHGGGNIPA- 476
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 436 KPGAFPPNGHVAGSGGL-----PPPPA 459
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 477 -----FVLGALIAAAVSGVLGTLVLPSPPPDA 502
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7
 G86360
 probable sucrose-proton symporter SUC2 protein T2ZJ18.12 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 02-Nov-2001
 C:Accession: G86360
 R:Thelodopsis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
 Chm, C.W.; Conway, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Cressy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
 A.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzali,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID: 21016719

A:Accession: G86360
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-512 <STO>
A:Cross-references: GB:AE005172; NID:g3287687; PIDN:AAC25515.1; GSPDB:GN001411
C:Genetics:
C:Map position: 1
C:Superfamily: common tobacco sucrose transport protein

| | | | | |
|-----------------------|------------------|--------------------|-------------|-------------|
| Query Match | 11.3%; | Score 323.5; | DB 2; | Length 512; |
| Best Local Similarity | 24.1%; | Pred. No. 6.7e-17; | | |
| Matches 123; | Conservative 92; | Mismatches 195; | Indels 101; | Gaps 18; |

| | | | |
|----|-----|---|-----|
| Qy | 14 | RKXLLLVNLTTEGLEUCIAGITVYRPLLEIEVGEKMTWVLIGPVILGVCPLLS | 73 |
| Db | 28 | RLRRTISVSSIAVGQGMALQSLTTPVYOLIGTTHKASLWLCGPTSGMDVPIYG | 87 |
| Qy | 74 | ASDHMRGRYRRRPFTHWLSLIGLLSLFLIPRAGWLAGL-----LCDPRELELALI | 126 |
| Db | 88 | HSDRCTSRFRRRRPFIVAGLVTVAFTL-----GYAADIGHSKGDOLDKPPKTRAIIFA | 144 |
| Qy | 127 | LGVLDDFCQGVCTPLEALLSDLFR-DPDHCRQASVYAFMSLGGCGY-----LL | 178 |
| Db | 145 | LGEWILDVANNLTLOGPCRAFLADLSAGNAKKTITTAALFFSFEWAQNVLTGAAGSTRILY | 204 |
| Qy | 179 | PAIDMD-TSALAPYLQEOECLFGLTLLIFLNCVANTLLVAEALGPTPEAGLSAPSL | 237 |
| Db | 205 | KVYFPTMTESODLYCANLKTCFELSTILL-LIYTFVSLCYVERK PPTPEPTADGKA--- | 259 |
| Qy | 238 | SPHCCPRARALAPNLGALLPRLHOLCCMRPRILRLFLPAELCSMMALMTFLFLFTDFVG | 297 |
| Db | 260 | -----SNVPF--FEETIGAFKEL-----KRPMMMLLTVALNTNLTAMPFLLEFDDIMWG | 305 |
| Qy | 298 | EGLYGVPRAPEPGEARHNDDEGVMSGLFLQCAISLVSFLVMDRLVORGETRAVYLA | 357 |
| Db | 306 | REYVGSNDATATAASKKLYINDGVBAIGALIMNALIYLGMSLGEVIMIGRKLS-GAKRLM | 364 |
| Qy | 358 | SVAAFPVAAGATCLISHVAV-----VTASA-----ALTG-----FT | 388 |
| Db | 365 | GIVFTL-----AICLAMPVVYTKOAEHHRDHGAKTGPPEGNVATGALLTLFALIGIPQAIT | 421 |
| Qy | 389 | FSALQILPYTLASLYHREK-----OVFLTPKYRGDVGAGASSEDLSMTSFLPGR | 435 |
| Db | 422 | FS-----IPFALASIFSTNSGAGGLSGVLTNALIAVPPQWIVSVGGPPELFCGGNIFA- | 476 |
| Qy | 436 | KPGAPFPNGHVGAGGSGLL-----PPPA | 459 |
| Db | 477 | -----FVLGATAAIVSGVLLATVLPSPPPA | 502 |

```

RESULT      8
S48789
sucrose transport protein - common tobacco
C:Species: Nicotiana tabacum (common tobacco)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 26-May-2000
C:Accession: S48789
R:Buerkle, X.Y.Z.: Frommer, W.B.
submitted to the EMBL Data Library, October 1994
A:Reference number: S48787
A:Accession: S48789
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-507 <BUE>
A:Cross-references: EMBL:X82276; NID:g575350; PID:g575351
;Superfamily: common tobacco sucrose transport protein

```

| | | | | |
|-----------------------|--------------|--|-----------------|-------------|
| Query Match | 11.1%; | Score 317.5; | DB 2; | Length 507; |
| Best Local Similarity | 24.3%; | Pred. No. 1.9e-16; | | |
| Matches 125; | Conservative | 87; | Mismatches 202; | Indels 99; |
| | | | | Gaps 18; |
| QY | 4 | RUMVSRLLRHKKAOILLVNLTFGIEVCLACGITTVPRLLEEGVEDEKFMIVLIGIPVL | 63 | |


```

Db 23 KLM-----KIMVASIAGVGFMAQLQLSLTFPVVQLGIPHKRASFIWLCGPS 72
OY 64 GLVCPPLIGASDHNRGRGRRPFIMWLSLGLSLFLIPRAGWLAGLCPDP-----R 118
Db 73 GMIVPVGIVGYSDNCSRRGRGFIAGALVTVIIFLIGFAADL-GHATGPDLKGSK 131
OY 119 PLELLALLIGVGLDCCGVCTPLEALLSDFRDPDHCROAVSYAFMISGGCGYLL 178
Db 132 PRAIAVGVGVFGLDVANMLOGPCRALADLSGKARRKTSNAFSPFMAVGNVIGYAA 191
OY 179 PAID-----WDSALAPYLQTOECLF-GLLTFLFCTVAATLLVAEEALGPTEPA 229
Db 192 GYSRLCKTFFPSKTPACDIYCANLKSCEFIIVFLLSLITLALT--VYRENELEPKDEH 249
OY 230 ECLSPASLSPPHCCPCRRARLAFRNGLALPRLHQLCCMRPTRLRLFVAILCSMALMTPT 289
Db 250 E-----IDKAGARKSKVPF--FGEIFGALKDL--PRPMWILLVTSIMNIARFPPE 297
OY 290 LEYTFVSGGLQGVPRAPRGTEARHVDGVRMSGLGLFQCAISLVFSYMDRLVQRF 349
Db 296 LVDTPMMAEVIYGG--KVGDG---RLYDLGVHAGALGILLNSVVGFSLSVEFLGKKI 351
OY 350 GTRAVYLVSAAPVPAAGATCLSHSVAVY-----TASALTGFTFSALQI-- 394
Db 352 G-GVRLMGLIFVL--AVCALVLYVTKMAEKSKQYAHGTIAPTSGVATIGALTFLA 407
OY 395 -----LPYTLASLYHREK-----QVFLPKYRGDTGASSEDLSMTS 430
Db 408 VLGIPLAVTFVYFPALASIFSSNAGSGGLSLGLVNLAIIVPQMLVSIAGPMDLFGGG 467
OY 431 FLGPKPGAPFPNGHVAGSGS-----LLPPPPA 459
Db 468 NLPG-----FTVGAVAAASGILATMLPSPPA 495

```

RESULT 9
S43142
sucrose transport protein - castor bean

N:Alternate names: sucrose carrier
C:Species: Ricinus communis (castor bean)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000
C:Accession: S43142
R:Weig, A.; Komor, E.
submitted to the EMBL Data Library, March 1994
A:Description: A sucrose carrier from Ricinus communis.
A:Reference number: S43142
A:Accession: S43142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <MEI>
A:Cross-references: EMBL:231561; NID:q468561; PID:q468562
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.9%; Score 311; DB 2; Length 533;
Best Local Similarity 23.3%; Pred. No. 6.2e-16;
Matches 134; Conservative 95; Mismatches 225; Indels 120; Gaps 19;

```

OY 17 LLLVNLTFGLFVCLAGITVPPPLLEVEGEEKMTWVLGIPVIGVCPPLIGSASD 76
Db 37 KYVMVASIAGIOPGMALQSLTFPVVQLGIPHTMAATWLCGPISGMLVPIVGYHSD 96
OY 77 HMRGVRGRRRPFIMWLSLGLSLFLIPRA--GWLAG-LLCPDPPELALLIGVGLL 132
Db 97 RCTSRFGRRRPFIMWLSLGLSLFLIPRA--GWLAG-LLCPDPPELALLIGVGLL 156
OY 133 DCCGVCTPLEALLSDF-RDPDHCROAVSYAFMISGGCGYLLPAID-----W 183
Db 157 DVANMLOGPCRALADLSGTSQKTRTANALFSPFMAVGNVIGYAAVYTHLYKLEPPT 216
OY 184 DYSALAPYLQTOECLFGLITLIFLTCVATLLVAEEALGPTEPAEG-----SAP 235

```

```

Db 217 KTTACDVYCANLKSCEFIISVILLSLTVLALSYKEK-----PMSPDQVADNEDDTASAQ 272
OY 236 SLSPHCCPCRRAL--AFRNLGALLPRLHQLCCMRPTRLRLYVAELCSMMALMTFTLFT 293
Db 273 SSASQMPPEFGEIGAFKNL-----KRPMMILLVTCMLNIAMPEFLIFT 318
OY 294 DVEGGLYGVPRAPRGTEARHVDGVRMSGLGLFQCAISLVFSYMDRLVQRFTRA 353
Db 319 DMNGREVIYGG--DSSASQQLKYDRGVAGALGLMLNSVVGFTSLGVEVLARGVG-GV 375
OY 354 VYLASVAAFVPAAGATCLSHSVAVVYVYASALTGFTFSALQIILPYTLASLYHREKQVFLPK 413
Db 376 KRLMGIVNFVL--AVCALMTV-LVTKQAEK-----RR 405
OY 414 YNGDTGASSEDLSMTSLFPGPKGAPFPNGHVAGSGGLLPPPPALCASACDVSVRY 473
Db 406 FATVSGAK-----VP-----SCVKKAGALFLAV 433
OY 474 VGEPTFARV-VP-----GRGICDLAIIIDSAFLISQVAPSLFMGSTVOL--SQS 519
Db 434 KGVPOAITYSIPPALASIFSNAGSGGLSLGLVNLAIIVPQMLVSIAGPMDLFGG 493
OY 520 VTAVYVSAAGLGLVATFATQVVFDDSLAKYSA 553
Db 494 LPAPVGAVALASGIFALTMLPSPPQDMPSAKA 527

```

RESULT 10
T14339

sucrose-proton transport protein - carrot
N:Alternate names: sucrose/H+ symporter protein

C:Species: Daucus carota (carrot)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C:Accession: T14339
R:Shakya, R.; Sturm, A.
Plant Physiol. 118, 1473-1480, 1998
A:Title: Characterization of source- and sink-specific sucrose/H+ symporters from car
A:Reference number: 217991; MUID:99063785
A:Accession: T14339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-501 <SHA>
A:Cross-references: EMBL:Y16766; NID:q2969886; PIDN:CAW76367.1; PID:q2969887
A:Experimental source: cultivar Nantaise; leaf
C:Genetics:
A:Gene: SUT1a
C:Superfamily: common tobacco sucrose transport protein

Query Match 10.7%; Score 305.5; DB 2; Length 501;
Best Local Similarity 21.6%; Pred. No. 1.5e-15;
Matches 124; Conservative 91; Mismatches 213; Indels 145; Gaps 14;

```

OY 18 LLLVNLTFGLFVCLAGITVPPPLLEVEGEEKMTWVLGIPVIGVCPPLIGSASD 77
Db 30 LIRVASVACGIOPGMALQSLTFPVVQLGIPHTMAATWLCGPISGMLVPIVGYHSD 89
OY 78 WBRGVRGRRRPFIMWLSLGLSLFLIPRAGWLAGL--CPDRPELALLIGVGLDPC 135
Db 90 CTSKGRRRRPFIMWLSLGLSLFLIPRAGWLAGL--CPDRPELALLIGVGLDPC 149
OY 136 GOVCTPLEALLSDF-RDPDHCROAVSYAFMISGGCGYLLPAIDMDTALAPYLGT 194
Db 150 NMWTOGPCRALADLSGTSQKTRTANALFSPFMAVGNVIGYAAVYTHLYKLEPPT 216
OY 195 OECEFLGLTLFLFNC-----VAATLLVAEEAL-----ATGAYSGW 199
Db 200 KYVFPFSLTSSCTINCANKSAFYIDITITITITISISAAERRRISQDQPSQSDCT 259
OY 233 SAPSLSPHCCPCRRARLAFRNGLALPRLHQLCCMRPTRLRLFVAILCSMALMTFTLFT 292
Db 260 AQ-----SGHI-----EAFIMELFGTFRLLPJGSVAVILLVTCMLMGMPFLFD 305

```

[illegible]

RESULT 11
S48788
sucrose transport protein - tomato (fragment)
C:Species: Lycopersicon esculentum (tomato)
C:Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 09-Jun-2000
C:Accession: S48788
R:Buerkle, X.Y.: Frommer, W.B.
submitted to the EMBL data library, October 1994
A:Reference number: S48787
A:Accession: S48788
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-428 <BUED>
A:Cross-references: EMBL:X82275
C:Superfamily: common tobacco sucrose transport protein

| | | | | | | | | | | |
|----|-----------------------|---|--------------|-------------|------------|--------|--------|-----|------|-----|
| | Query Match | 10.4% | Score | 297.5 | DB 2: | Length | 428: | | | |
| | Best Local Similarity | 25.8%: | Pred. | No.5.2e-15: | | | | | | |
| | Matches | 114: | Conservative | 82: | Mismatches | 163: | Indels | 83: | Gaps | 18: |
| OY | 4 | RLLVSRFLRRHRAQQLLVNLLTFGEVCVCAAIITVPRLLEVEGEEKEMTVLIGIPVL | 63 | | | | | | | |
| Db | 25 | KLM-----KTIIVASIAAGVFGMALQSLTLTPYOLLGIRPHRASFTWLOGPIS | 74 | | | | | | | |
| OY | 64 | GLVGVPLLGASADHWGRGRRRPRTIWAISLTGLSTFLPIRAGWLALGPCDP-----R | 118 | | | | | | | |
| Db | 75 | GMIQGVNVGVYSNDSCSSRGRRPRPTAAGAALVTIATVFLTGPAAALD-GHASGDPLKGSK | 133 | | | | | | | |
| OY | 119 | PLEALLILGLVGILLDDCGOYCFFPLEBALLSDLFRRPD-HCRQAYSYTAAMISLGGCTG- | 176 | | | | | | | |
| Db | 134 | PRAIAAVGVFWLLIDVANNNMLQGPPCALLALDSGKSGMKRMANAFSEFFKAGNTLLGA | 193 | | | | | | | |
| OY | 177 | -----LRLALDMD-TSALAPYLGPQECLLF-GLTLRFIFETCVAAATL-----VAEEAAG | 224 | | | | | | | |
| Db | 194 | AGSYSLRFKVFPPSKTKACDMTRANKSCFFTAIFLILTLTIALTLVENELPEKEELE | 253 | | | | | | | |
| OY | 225 | PTPAEGLSAPLSLPCCPCPARLAERNLGALLPRLHOLCRRMPTRLRLFAVELCSMAA | 284 | | | | | | | |
| Db | 254 | IDEKLSGAG-----KSKVPP-FGEIENGALKDL-----PRRMWILLTYTCMLNTA | 296 | | | | | | | |
| OY | 285 | LMTFTLLFYTDVEGELYOGVPRAEPTGARRHRYDEGVRMGSGLFLOCAISLVSFLVMDR | 344 | | | | | | | |
| Db | 297 | WPFPEFLYDFDMAKEEVFG--QVGD-AKTVDGLVRGAGALGLLIQSVALGPMISGVFEF | 350 | | | | | | | |
| OY | 345 | LVOFGFRANYYLASVAAPYAAGATLSHSVAVY-----TASAALTGF----- | 388 | | | | | | | |
| Db | 351 | IGKKIG-GAKRIIGIINFLV---AICIAMTIIWTKNKAERSRGHDAGITGMPPGYKIGA | 406 | | | | | | | |
| OY | 389 | ---ESALQI-----LPTYLA | 400 | | | | | | | |
| Db | 407 | LLEFAALGIPLAVTFSIPEALA | 428 | | | | | | | |

RESULT 12
S38657
sucrose transport protein ptpl - common plantain
N:Alternate names: sucrose transporter ptpl
C:Species: Plantago major (common plantain)
C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 26-May-2000
C:Accession: S38657
R:Garritz, M., Stolz, J.; Sauer, N.
Submitted to the EMBL Data Library, November 1993
A:Reference number: S38657
A:Accession: S38657
A:Molecule type: mRNA
A:Residues: 1-510 (GAH>
A:Cross-references: EMBL:X75764; NID:g415987; PIDN:CAAS3390.1; PID:g415988
C:Genetics:
A:Gene: ptpl
C:Superfamily: common tobacco sucrose transport protein
C:Keywords: sugar transport

| | | | | |
|-----------------------|------------------|--------------------|-------------|-------------|
| Query Match | 10.3%; | Score 294; | DB 2; | Length 510; |
| Best Local Similarity | 24.7%; | Pred. No. 1.1e-14; | | |
| Matches 126; | Conservative 77; | Mismatches 199; | Indels 108; | Gaps 19; |

```
QY    18 LLVNLFFLEGYCLAGITVPVRLLEVEVEEKKEMPMUJGTPVLGVLCPLTGSASH   77
      :|::||::||::||::||::||::||::||::||::||::||::||::||:
Db    29 IFLVAIATAGVOFCMAIOLSLILTPYOOLLGIPIHKMASYIMLCPISGMIVQPVVGFSDN   88
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    78 WKGRYRRRPETMALSGILT-LSLELI PRAGWLGL-----LCDPRPLEALLILCY   129
      |||||||::||::||::||::||::||::||::||::||::||::||:
Db    89 CTSRFGRRRPET-AAGAGLVAVAAVLII---GFAADLGHAIGDSICDGLKPRAI GVEVFEG   144
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    130 GLIDPCGGOVCTPBLEALISDLF-RDPHCQRQAYSVAFMUSLGGCGLCYLLPAID-----   182
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    145 WLIDVANNNMLOGPCRALLADI SGNVTKKMANANSFEFFMAVGNVLTGAAGSISRMTKYE   204
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    183 ---WDYSALAPYLGTQOEC-LFGLTLTFLTVCVAATLL-----VAEHALGPTEPBGLSA   234
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    205 PPSKTAKADIYCANLKSCFIISTILFTLTTLATISIYREKRHNAEEBYY--TAACKGFKI   261
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    235 PSLSHHCCRPCRARLFNRULGMLLRHLQLCOCRMRTLRLEFVALCEGSMALMFTELYTD   294
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    262 P-----VPBELGALKDLPREPMMILLVTALNMWIMEGELFLFDTD   301
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    295 FVGEGLYGVBPRAEPBTERRRHDEGVNMSGLSGLELOCAISLVEFSVMDBRVORFC----   350
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    302 WMGREVY-GEQQHKAPELAITYNKVSAGALSILMSIYLGFASTLGOVYMARALGVKR   360
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    351 -----TRAYILA-SVAFPYAAGATCLSNAVYNASALNG-FTFESALO-----   394
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    361 IMGVNVELIATCLOKTIYTIRVASHHRPYSNGVLOTPEBSVKIGALVYESALIGLIAITF   420
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    395 -LPYTLASYIHREKOVLPKRYKGDTGASSEDSTL-----MTSEFLPGBKPG--   438
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    421 SVPPALASTY-----STYTGSGQGSLSIGVLNALIVIPQMIVVASA GSPDMIFG   468
      :|::||::||::||::||::||::||::||::||::||::||::||:
QY    439 ----APPNGNHAGACSG-----LLPRPPA   459
      :|::||::||::||::||::||::||::||::||::||::||::||:
Db    469 GGNDLPASFVGAATAAASGIFAETMI SPSPA   498
```

RESULT 13
S5114
sucrose-proton symporter - beet
C:Species: *Beta vulgaris* (beet)
C:Date: 07-May-1995 #sequence_revision 01-Sep-1995 #text_change 26-May-2000
C:Accession: S5114
R:Westram, A.; Eckhardt, U.; Frommer, W.B.; Riesenleier, J.W.
submitted to the EMBL data library, January 1995
A:Description: Sequence of a sugar beet sucrose transporter cDNA.
A:Reference number: S5114

A:Accession: S51114
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-523 <EMBL>
A:Cross-references: EMBL:X83850; NID:g633171; PID:g633172
C:Superfamily: common tobacco sucrose transport protein

| | | | | |
|-----------------------|------------------|--------------------|-------------|-------------|
| Query Match | 10.3% | Score 294 | DB 2 | length 523; |
| Best Local Similarity | 24.3% | Pred. No. 1.2e-14; | | |
| Matches 125; | Conservative 78; | Mismatches 199; | Indels 112; | Gaps 20. |

| | | | |
|----|-----|---|-----|
| QY | 17 | OLLVNLTFTEGVLNIAITFYVPRLLLEVGEKEMTWLIGIPVLGICVPLTASGD | 76 |
| Db | 38 | KLALVSLAAGVQFMALQQLSLTPYVQGLGIPTHTAPYIWLICGPISGMIVQTPVGYSD | 97 |
| QY | 77 | HWRCYGRRRPPTMALSLGILLSLFLIPAGMLA--GLLCPD-----PRPELALLTLGV | 129 |
| Db | 98 | RCTSFGRGRSPRI--AVGATLVGEFVNSLIGFADIGHMTGDPNGVMPKREAIATVEVYGF | 154 |
| QY | 130 | GLLDRCGVVCFPLEALLSLDLPDPD--HCRQAVSVYAFMISLGCGLGY-----LLP | 179 |
| Db | 155 | WILDVANNTLGGPCALLADMAAGSOAKTRYANAFESFEMALDNIGCYAGSYGRLYVPR | 214 |
| QY | 180 | AIDMDFSLAPRYLQTEBCELFGLTLPLFLTCVAATLLVAEEALGTPBAEGLSAPLS | 239 |
| Db | 215 | FT--HTKACDPTCYAMIKSCOFFISITLLIVLTILIALSVRRER----PPTLDEIOEEEMLKN | 268 |
| QY | 240 | HCCPRRLARFANLICALLLPRLHLQCCRMPTLRFLVFAELCSMAMLTFTLFTDPTGEG | 299 |
| Db | 269 | NTGGC--ALLPR--FQQLGALKDL----PKPMLILLVLCIMIANPFLFLPTDMWGRE | 321 |
| QY | 300 | LYQGVRAEPGTEARRHYDEGVRMGSLGFLQCAISLVESLVMDRLVQVRFG--TRAVYLLA | 357 |
| Db | 322 | YVGGT-----VGEKADMGVNAAGALGIMINSVVLGIMSLGIEKTLARLVGGVKRLMGIV | 375 |
| QY | 358 | SVAAPFVAAGATCTLSHVAVVTASA-----ALTG | 387 |
| Db | 376 | NLIL-----AVCLAMTI--LVTKSABHRYATHNHVPGALGPRLPRPGVKGALAIFFALGI | 428 |
| QY | 388 | TFESALQILPYTLASLYHREK-----OVFLPKRYGDTGASSEDLSMT | 429 |
| Db | 429 | PLATFSIPFALASIFSSAGSGQGLSLGVNLDAIVVQPMFVSVTSGPW-----DAL-- | 480 |
| QY | 430 | SFLPRKPGAPRRPNHVGAGSG-----LLPRP | 458 |
| Db | 481 | -FEGGNLPA--FVVGAVAAATASALISFTLLPP | 511 |

RESULT 14
 F96741
 Probable sucrose transport protein F17M19.4 [imported] - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: F96741
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso
 Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
 ansen, N.F.; Hughes, B.; Hultzart, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, O.
 C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maitl, R.; Marziani,
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
 A:Reference number: A86141, MUID:21016719
 A:Accession: F96741
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-512 <SRO>
 A:Cross-references: GB:AE005173; NID:96978914; PIDDN:AAF3406.1; GSPDB:GN00141
 ;Genetics:
 ;Gene: F17M19.4

A;Map position: 1
C;Superfamily: common tobacco sucrose transport protein

| | | | | |
|---------------------------|-------|-------------------|------------|------------|
| Query Match | 10.2% | Score 292.5 | DB 2 | Length 512 |
| Best Local Similarity | 23.8% | Pred. No. 1.5e-14 | | |
| Matches 120; Conservative | 87; | Mismatches 211; | Indels 87; | Gaps 18; |

[illegible]

RESULT 15
T02982
Probable sucrose transport protein - rice
C:Species: *Oryza sativa* (rice)
C:Date: 24-Mar-1999 #sequence.revision 24-Mar-1999 #text.change 21-Jul-2000
C:Accession: T02982
R:Hirose, T.; Imazumi, N.; Scofield, G.N.; Furukawa, R.T.; Onogi, R.
Plant Cell Physiol. 38, 1389-1396, 1997
A:Title: cDNA cloning and tissue specific expression of a gene for sucrose transporter
A:Reference number: Z14809; MUID:98182940
A:Accession: T02982
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-537 <HIR>
A:Cross-references: EMBL:D87819; NID:g2723470; PTDN:BA24071.1; PID:g2723471
A:Experimental source: cultivar Nipponbare, leaf
C:Genetics:
A:Gene: SUT1
A:Superfamily: common tobacco sucrose transport protein

| | | | | | |
|-----|---|------------------|-----------------|------------|-------------|
| | Query Match | 10.2% | Score 292; | DB 2; | Length 537; |
| | Best Local Similarity | 26.7%; | Pred. No. | 1.7e-14; | |
| | Matches 115; | Conservative 72; | Mismatches 167; | Indels 76; | Gaps 19; |
| QY | 17 QLLLVNLTFLGLEVCCLAAGITVVPPLLEVGVEEKFMVMVLGIGIVGLACVCPILGSASD | 76 | | | |
| | : : : : : : : : : : : : : : : : | | | | |
| bob | 51 RLIIISGVAVGGVQQGMALQLSLTPRVPTLISHALTFSFMMICGSIADMVOPPCGVATSD | 110 | | | |

```

QY 77 HMRGRYGRRPFIWALSIGILLSLELI--PRAGWLAGLLCPD-----PPLEIALLIL 127
Db 111 RCTSKMGRRRRPYILTGCVLICLAVVIGFSADIGYAMGDTKEDCSVYHGRHAAIYVVL 170
QY 128 GVGILLDFGQVCFTPLEALLSDF--RDPDHCRAQSYVYAFMISLGCGTGLPAID--- 182
Db 171 GFWLLDFSNNTVQGPAPALMADLSGRHGP--TANSTFCGMMAMGNILGYSSGSTNNMH 227
QY 183 -WDTSALAPYLGTQEBEC-----LFG--LTLILFLT-CVAATLLVAEE-----AALGPT- 226
Db 228 KM-----PFLKTRACCACANLKGAFIVAYIFLSICIVITLIFAKEVPEFGNAL-PTK 281
QY 227 --EPAEGLSAPSLSPHCCPCRARLAFRNLGALLPRLHOLCCRMPTLRRLFVAELCSMA 284
Db 282 SNEPAEPEGTG-----PLAVLKGFRLPTGMPSV-----LIVTGL-TWLS 320
QY 285 LMTFTLEFYTDVFGELVYQGVPR-AEPGTEARRHYDEGRMGSIGIFLOCAISIVFSLYMD 343
Db 321 WPPFLYDTDMGRIYHGDPKGTDPQIEA--FNQGVRAAGAFGLLINSIVLAGFSSFLIE 377
QY 344 RLVOFGTRAVYLASVAAPVAAGATCL-----SHVA--VVTASALTGFTFSALQ 393
Db 378 PMCKRVGPRVYVWTSNPLVCIAMAAATALISFWSLKDFFHTVQKATTADKSIKAVCLVLEA 437
QY 394 ILPYTLASLY 403
Db 438 FLGVPLAVLY 447

```

Search completed: June 26, 2002, 13:45:46
 Job time: 3939 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:49:02 : Search time 30.25 Seconds
(without alignments)
102.812 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28

Sequence: 1 DFVGEGLYGVPRAGTEARRHVDEGVR 28

Scoring table:

OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 268091

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Listing first 45 summaries

Database :

A-Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 13 | 46.4 | 25 | 22 | AAU69815 |
| 2 | 13 | 46.4 | 25 | 22 | AAU01170 |
| 3 | 13 | 46.4 | 25 | 22 | AAU69815 |
| 4 | 13 | 46.4 | 25 | 20 | AAU69815 |
| 5 | 13 | 46.4 | 27 | 21 | AAU69815 |
| 6 | 13 | 46.4 | 27 | 22 | AAU69815 |
| 7 | 13 | 46.4 | 27 | 22 | AAU04208 |
| 8 | 13 | 46.4 | 27 | 22 | AAU01296 |
| 9 | 11 | 39.3 | 15 | 22 | AAU69813 |
| 10 | 11 | 39.3 | 15 | 22 | AAU01168 |
| 11 | 11 | 39.3 | 15 | 22 | AAU69813 |

| | | | | | |
|----|------|----|----|----------|--------------------|
| 12 | 21.4 | 17 | 22 | AAU69814 | Human prostate pro |
| 13 | 21.4 | 17 | 22 | AAU01169 | Human prostate-spe |
| 14 | 21.4 | 17 | 22 | AAU69815 | Human prostate-spe |
| 15 | 17.9 | 8 | 22 | AAU01151 | Peptide #1 of pep |
| 16 | 17.9 | 9 | 20 | AAU21072 | Human p53 cellular |
| 17 | 17.9 | 9 | 20 | AAU21072 | M. spicata geranyl |
| 18 | 17.9 | 9 | 22 | AAU69813 | HER2/NEU DR super |
| 19 | 17.9 | 10 | 22 | AAU69813 | Human complement |
| 20 | 17.9 | 13 | 17 | AAU69813 | HER2/NEU DR super |
| 21 | 17.9 | 14 | 17 | AAU69813 | HER2/NEU DR super |
| 22 | 17.9 | 14 | 17 | AAU69813 | HER2/NEU DR super |
| 23 | 17.9 | 14 | 19 | AAU69813 | HER2/NEU DR super |
| 24 | 17.9 | 15 | 12 | AAU69813 | HER2/NEU DR super |
| 25 | 17.9 | 15 | 14 | AAU69813 | HER2/NEU DR super |
| 26 | 17.9 | 15 | 19 | AAU69813 | HER2/NEU DR super |
| 27 | 17.9 | 15 | 22 | AAU69813 | HER2/NEU DR super |
| 28 | 17.9 | 17 | 15 | AAU69813 | HER2/NEU DR super |
| 29 | 17.9 | 18 | 16 | AAU69813 | HER2/NEU DR super |
| 30 | 17.9 | 18 | 21 | AAU69813 | HER2/NEU DR super |
| 31 | 17.9 | 19 | 18 | AAU69813 | HER2/NEU DR super |
| 32 | 17.9 | 19 | 20 | AAU69813 | HER2/NEU DR super |
| 33 | 17.9 | 20 | 16 | AAU69813 | HER2/NEU DR super |
| 34 | 17.9 | 21 | 17 | AAU69813 | HER2/NEU DR super |
| 35 | 17.9 | 21 | 19 | AAU69813 | HER2/NEU DR super |
| 36 | 17.9 | 22 | 19 | AAU69813 | HER2/NEU DR super |
| 37 | 17.9 | 22 | 22 | AAU69813 | HER2/NEU DR super |
| 38 | 17.9 | 26 | 21 | AAU69813 | HER2/NEU DR super |
| 39 | 17.9 | 26 | 21 | AAU69813 | HER2/NEU DR super |
| 40 | 17.9 | 26 | 21 | AAU69813 | HER2/NEU DR super |
| 41 | 17.9 | 26 | 21 | AAU69813 | HER2/NEU DR super |
| 42 | 17.9 | 27 | 22 | AAU69813 | HER2/NEU DR super |
| 43 | 17.9 | 27 | 22 | AAU69813 | HER2/NEU DR super |
| 44 | 14.3 | 4 | 18 | AAU69813 | HER2/NEU DR super |
| 45 | 14.3 | 5 | 9 | AAU69813 | HER2/NEU DR super |

ALIGNMENTS

| | | |
|----------|---|---------------------------|
| RESULT 1 | AAU69815 | standard; Peptide: 25 AA. |
| ID | AAU69815 | |
| XX | AAU69815 | |
| AC | AAU69815 | |
| XX | AAU69815 | |
| DT | 30-JAN-2002 | (first entry) |
| XX | 30-JAN-2002 | |
| DE | Human prostate protein P703P antigenic epitope #12. | |
| XX | Human prostate protein P703P antigenic epitope #12. | |
| KW | Human; prostate cancer; ss; cytosolic; immunostimulant; tumour; | |
| KW | antigen; epitope. | |
| XX | | |
| OS | Homo sapiens. | |
| XX | | |
| PN | WO200173032-A2. | |
| XX | | |
| PD | 04-OCT-2001. | |
| XX | | |
| PF | 27-MAR-2001; 2001WO-US09919. | |
| XX | | |
| PR | 27-MAR-2000; 2000US-0536857. | |
| PR | 09-MAY-2000; 2000US-0568100. | |
| PR | 12-MAY-2000; 2000US-0570737. | |
| PR | 13-JUN-2000; 2000US-0593793. | |
| PR | 27-JUN-2000; 2000US-0605783. | |
| PR | 10-AUG-2000; 2000US-0636215. | |
| PR | 29-AUG-2000; 2000US-0651236. | |
| PR | 06-SEP-2000; 2000US-0652729. | |
| PR | 02-OCT-2000; 2000US-0679426. | |
| PR | 10-OCT-2000; 2000US-0685166. | |
| XX | | |
| PA | (CORI-) CORIXA CORP. | |
| XX | | |

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAW, Hepler WT, Henderson RA;
 XX
 DR WPI: 2001-639232/73.
 PT New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 PS Claim 2; Page 403; 579pp; English.
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC peptide of the invention. The peptides either represent antigenic
 CC epitopes or domains of prostate specific proteins.
 CC
 SQ Sequence 25 AA;
 XX
 Query Match 46.4%; Score 13; DB 22; Length 25;
 Best Local Similarity 100.0%; Pred. NO. 6e-06; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0;
 QY 3 VGEGLYGVPRAE 15
 ||||||||||||
 Db 1 vgeglygvprae 13
 XX
 RESULT 2
 AAM01170
 ID AAM01170 standard; Peptide: 25 AA.
 XX AAM01170;
 AC
 XX 04-OCT-2001 (first entry)
 DT
 XX Human prostate-specific peptide used in epitope mapping studies #27.
 DE
 XX Human: prostate cancer; prostate-specific; diagnosis; vaccine;
 KW cytostatic; gene therapy; metastasis.
 KM
 XX Homo sapiens.
 OS
 XX WO200151633-A2.
 PN
 XX 19-JUL-2001.
 PD
 XX 16-JAN-2001; 2001WO-US01574.
 PF
 XX 14-JAN-2000; 2000US-0483672.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
 PI Wang A, Meagher MJ;
 PL
 DR WPI: 2001-425873/45.
 XX
 DR New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines -
 XX

PS Claim 2; Page 403; 543pp; English.
 XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II), (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
 CC AAM01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 CC
 SQ Sequence 25 AA;
 XX
 Query Match 46.4%; Score 13; DB 22; Length 25;
 Best Local Similarity 100.0%; Pred. NO. 6e-06; Indels 0; Gaps 0;
 Matches 13; Conservative 0; Mismatches 0;
 QY 3 VGEGLYGVPRAE 15
 ||||||||||||
 Db 1 vgeglygvprae 13
 XX
 RESULT 3
 AAG99055
 ID AAG99055 standard; Peptide: 25 AA.
 XX AAG99055;
 AC
 XX 25-SEP-2001 (first entry)
 DT
 XX Human prostate-specific epitope mapping peptide #27.
 DE
 XX Human: prostate cancer; therapy; diagnosis; cat eye syndrome;
 KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
 KM prostate specific antigen; PSA.
 KW
 XX Homo sapiens.
 OS
 XX WO200134802-A2.
 PN
 XX 17-MAY-2001.
 PD
 XX 09-NOV-2000; 2000WO-US30904.
 PF
 XX 12-NOV-1999; 99US-0439313.
 PR
 XX 18-NOV-1999; 99US-0443686.
 PR
 XX (CORI-) CORIXA CORP.
 PA
 XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
 PI
 DR WPI: 2001-308785/32.
 XX
 DR Isolated polypeptide comprising at least an immunogenic portion of a
 PT prostate-specific protein, useful in the diagnosis and therapy of a
 PT prostate cancer -
 XX
 PS Claim 3; Page 303; 325pp; English.
 CC The present invention describes an isolated polypeptide (PI) comprising
 CC at least an immunogenic portion of a prostate-specific protein, or its
 CC variant. Also described are polynucleotides (NI) encoding (PI). (PI) and
 CC (NI) have cytostatic activity and can be used in vaccine production.
 CC The polypeptides, nucleic acids and antibodies from the present
 CC invention are useful in the diagnosis and therapy of prostate cancer.

CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
 CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
 CC region. Prostate specific antigen (PSA) P501S was located on
 CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
 CC polynucleotide and polypeptide sequences used in the exemplification
 CC of the present invention.

XX Sequence 25 AA;

Query Match 46.4%; Score 13; DB 22; Length 25;
 Best Local Similarity 100.0%; Pred. No. 6e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VEGGLYQGVRAE 15
 |||||
 Db 1 yegglyqvgrae 13

RESULT 4

ID AAW85071 standard; Peptide: 27 AA.

AC AAW85071;

XX 12-FEB-1999 (first entry)

DE PS108 protein derived peptide.

XX PS108 gene; prostate disease; benign prostatic hyperplasia; BPH;

KW prostatitis; prostatic intraepithelial neoplasia; PIN; cancer;

KW drug screening; gene therapy.

XX Homo sapiens.

PN WO9850567-A1.

XX 12-NOV-1998.

PF 01-MAY-1998; 98WO-US08930.

XX 02-MAY-1997; 97US-0850713.

XX (ABBO) ABBOTT LAB.

PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN;

PI Gordon J, Granados EN, Hodges SC, Klaas MR, Kratochvil JD;

PI Roberts-Rapp L, Russell JC, Stroupe SD;

DR WPI: 1999-034731/03.

XX New isolated prostate-specific polynucleotides - used to develop

PT products for the diagnosis and treatment of prostate diseases, e.g.

PT benign hyperplasia, prostatic or prostate cancer

XX Claim 17; Page 101; 122pp; English.

XX The present sequence is derived from the PS108 protein. The

CC specification describes a method for detecting the presence

CC of a target PS108 polynucleotide in a test sample. The method

CC comprises contacting the test sample with at least 1 PS108-specific

CC polynucleotide or complement, and detecting the presence of the target

CC PS108 polynucleotide. The products can be used for detecting, diagnosing,

CC staging, monitoring, prognosticating, in vivo imaging, preventing or

CC treating, or determining predisposition to diseases or conditions of

CC the prostate such as benign prostatic hyperplasia (BPH), prostatitis,

CC prostatic intraepithelial neoplasia (PIN) and cancer. In particular the

CC products can be used in drug screening and gene therapy.

XX Sequence 27 AA;

XX Query Match 46.4%; Score 13; DB 20; Length 27;

Best Local Similarity 100.0%; Pred. No. 6.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEARRHYDEGVR 28
 |||||
 Db 12 gtearrhydegvr 24

RESULT 5

ID AAB29271 standard; Protein: 27 AA.

AC AAB29271;

XX 07-FEB-2001 (first entry)

DE Human prostate-related PS108 partial protein sequence SEQ ID NO: 39.

XX Human; prostate cancer; PS108; antibody; tumour; metastasis.

XX Homo sapiens.

PN US6130043-A.

XX 10-OCT-2000.

PF 01-MAY-1998; 98US-0071710.

XX 02-MAY-1997; 97US-0850713.

XX (ABBO) ABBOTT LAB.

PI Cohen M, Colpitts TL, Friedman PN, Gordon J, Granados EN;

PI Billing-Medel PA, Klaas MR, Roberts-Rapp L, Stroupe SD, Yu H;

PI Kratochvil JD, Russell JC, Hodges SC;

DR WPI: 2000-655655/63.

XX Methods for detecting target prostate-specific polynucleotides or

PT diseases of the prostate (e.g. prostate cancer), comprising detecting

PT the presence of any of PS108 nucleic acid sequences in a test sample -

XX Example 10; Column 87-88; 55pp; English.

XX The present invention is related to a number of partial coding and

CC protein sequences for the human prostate tissue protein PS108. These

CC sequences can be used in the diagnosis and prognosis of prostate

CC diseases, particularly prostate cancer. They can also be used to produce

CC antibodies which can be used in treatment. The present sequence is one

CC of the PS108 partial protein sequences.

XX Sequence 27 AA;

OY 16 GTEARRHYDEGVR 28
 |||||

Db 12 gtearrhydegvr 24

RESULT 6

ID AAU69941 standard; Peptide: 27 AA.

AC AAU69941;

XX 30-JAN-2002 (first entry)

DE Human prostate protein P501S peptide 296-322.

XX Query Match 46.4%; Score 13; DB 21; Length 27;

KW Human; prostate cancer; ss; cytostatic; immunostimulant; tumour;
KW antigen; epitope.
XX
OS Homo sapiens.
FN WO200173032-A2.
PD
PD 04-OCT-2001.
PF 27-MAR-2001; 2001WO-US09919.
XX
XX 27-MAR-2000; 2000US-0536857.
PR 09-MAY-2000; 2000US-0568100.
PR 12-MAY-2000; 2000US-0570737.
PR 13-JUN-2000; 2000US-0593793.
PR 27-JUN-2000; 2000US-0605783.
PR 10-AUG-2000; 2000US-0636215.
PR 29-AUG-2000; 2000US-0651236.
PR 06-SEP-2000; 2000US-0657279.
PR 02-OCT-2000; 2000US-0679426.
PR 10-OCT-2000; 2000US-0685166.
XX
XX (CORI-) CORIXA CORP.
PA
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
XX WPI: 2001-639232/73.
XX
XX New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer -
XX
XX Claim 2; Page 558; 579pp; English.
XX
XX The invention relates to isolated prostate-specific
CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of
CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC the antigen-presenting cells are useful for stimulating and/or expanding
CC T cells specific for a tumour protein, and for inhibiting the development
CC of cancer especially prostate cancer. Compositions comprising the
CC polynucleotide and/or polypeptide are useful for stimulating an immune
CC response, and for treating cancer. The oligonucleotide is useful for
CC detecting cancer. The present sequence is a prostate specific
CC peptide of the invention. The peptides either represent antigenic
CC epitopes or domains of prostate specific proteins.
CC
XX
SQ Sequence 27 AA:
OY Query Match 46.4%; Score 13; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 3 VEGGLYGVPRAE 15
1 Vgegllygvprae 13
RESULT 7
AAU04208
ID AAU04208 standard; Peptide: 27 AA.
XX
XX AAU04208;
AC
XX
XX 23-OCT-2001 (first entry)
DT
XX Prostate-specific polypeptide PS108 #4.
DE
XX Prostate; PS108; Immunogen; drug screening; image localisation;
KW

KW diagnostic; therapeutic; prostate tissue disease; cancer; metastasis;
KW expressed sequence tag; EST.
XX
OS Homo sapiens.
FN US6252047-B1.
PD
PD 26-JUN-2001.
PF 15-MAR-2000; 2000US-0525397.
XX
XX 01-MAY-1998; 98US-0071710.
PR 02-MAY-1997; 97US-0850713.
XX
XX (ABBO) ABBOTT LAB.
PA
PI Billing-Medel PA, Cohen M, Colpitts TL, Friedman PN, Gordon J;
PI Granados EN, Hodges SC, Kias MR, Kratochvil JD, Roberts-Rapp L;
PI Russell JC, Stroupe SD, Yu H;
XX Russell JC, Stroupe SD, Yu H;
XX WPI: 2001-424488/45.
XX
XX Novel PS108 polypeptide useful in assays for detecting antibodies to
PT prostate tissue, and as immunogens to produce PS108 antibodies -
PT
XX Claim 1; Column 87; 55pp; English.
XX
XX The sequence represents the amino acid sequence of prostate-specific
CC PS108 #4. The PS108 polypeptide is useful in assays for detecting
CC antibodies to prostate tissue, and as an immunogen to produce antibodies.
CC The polypeptide is useful for screening compounds which specifically bind
CC to the polypeptide and for screening for drugs, compounds, or any other
CC agent which can be used to treat diseases associated with PS108. The
CC antibody is useful to detect, or for image localisation of, PS108 antigen
CC in a patient, for detecting or diagnosing a disease or condition, as
CC delivery agents for therapeutic agents as well as for diagnostic tests
CC and for screening for diseases or conditions associated with PS108,
CC especially cancer. The antibody is also useful for generating chimeric
CC antibodies for therapeutic use, for inhibiting the biological activity of
CC PS108, in therapy (for e.g. to treat prostate tissue disease including
CC prostate cancer and its metastases), and to detect the presence of any
CC determinants with the PS108 polypeptide.
CC
XX
SQ Sequence 27 AA:
OY Query Match 46.4%; Score 13; DB 22; Length 27;
Best Local Similarity 100.0%; Pred. No. 6.4e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 16 GTEARRHYDEGVR 28
12 gtearrhydegvr 24
RESULT 8
AAM01296
ID AAM01296 standard; Peptide: 27 AA.
XX
XX AAM01296;
AC
XX
XX 04-OCT-2001 (first entry)
DT
XX P501S peptide 296-322.
DE
XX Human; prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytotatic; gene therapy; metastasis.
XX
XX Homo sapiens.
OS
XX W0200151633-A2.
PN
XX

PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001WO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAM;
 PI Wang A, Meagher MJ;
 XX
 DR WPI; 2001-425873/45.
 XX
 PT New polynucleotide encoding a prostate-specific protein, for
 PT diagnosing, monitoring and treating prostate cancer in a patient and
 PT for use in vaccines -
 XX
 PS
 PS Claim 2; Page 522; 543pp; English.
 XX
 CC The present invention describes polynucleotide sequences (I) which encode
 CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
 CC and can be used in vaccine production and gene therapy. (I), (II),
 CC antibodies to (II), fusion proteins comprising (II), and isolated
 CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
 CC (I) and the antibodies are also used in the detection of cancer in a
 CC patient. The cancer that is diagnosed or treated is particularly
 CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
 CC (I) can be used for monitoring the progression of cancer in a patient.
 CC (I) and (II) can also be used to improve diagnostic and therapeutic
 CC methods for prostate cancer. They can indicate the level of metastasis
 CC as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
 CC AA01318 represent polynucleotide and amino acid sequences used in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 27 AA;

Query Match 46.4%; Score 13; DB 22; Length 27;
 Best Local Similarity 100.0%; Pred. No. 6.4e-06;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VEGGLYQGVRAE 15
 |||||
 DB 1 vgeglyggyprae 13

RESULT 9
 AA069813
 ID AA069813 standard; Peptide: 15 AA.
 XX
 AC AA069813;
 XX
 DT 30-JAN-2002 (first entry)
 XX
 DE Human prostate protein P703P antigenic epitope #10.
 XX
 KM Human; prostate cancer; ss: cytostatic; immunostimulant; tumour;
 KM antigen; epitope.
 XX
 OS Homo sapiens.
 OS
 PN WO200173032-A2.
 PD
 PD 04-OCT-2001.
 PF
 PF 27-MAR-2001; 2001WO-US09919.
 XX
 PR 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.

PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0657279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
 XX
 DR WPI; 2001-639232/73.
 XX
 PT New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 PS
 PS Example 19; Page 403; 579pp; English.
 XX
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumor protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC peptide of the invention. The peptides either represent antigenic
 CC epitopes or domains of prostate specific proteins.
 XX
 SQ Sequence 15 AA;

Query Match 39.3%; Score 11; DB 22; Length 15;
 Best Local Similarity 100.0%; Pred. No. 0.00034;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARRHYDEG 26
 |||||
 DB 5 gtearrhydeg 15

RESULT 10
 AAM01168
 ID AAM01168 standard; Peptide: 15 AA.
 XX
 AC AAM01168;
 XX
 DT 04-OCT-2001 (first entry)
 XX
 DE Human prostate-specific peptide used in epitope mapping studies #25.
 XX
 KM Human; prostate cancer; prostate-specific; diagnosis; vaccine;
 KM cytostatic; gene therapy; metastasis.
 XX
 OS Homo sapiens.
 OS
 PN WO200151633-A2.
 PD
 PD 19-JUL-2001.
 PF
 PF 16-JAN-2001; 2001WO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
 PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAM;

| | |
|----|---|
| PI | Wang A, Meagher MJ: |
| XX | |
| XX | WPI, 2001-425873/45. |
| DR | |
| XX | |
| PT | New polynucleotide encoding a prostate-specific protein, for |
| PT | diagnosing, monitoring and treating prostate cancer in a patient and |
| PT | for use in vaccines - |
| XX | |
| XX | |
| PS | Example 19; Page 402; 543pp; English. |
| XX | |
| CC | The present invention describes polynucleotide sequences (I) which encode |
| CC | prostate-specific proteins (II). (I) and (II) have cytostatic activity, |
| CC | and can be used in vaccine production and gene therapy. (I), (II), |
| CC | antibodies to (II), fusion proteins comprising (II), and isolated |
| CC | T cells prepared using (I) or (II) are used to treat cancer in a patient. |
| CC | (I) and the antibodies are also used in the detection of cancer in a |
| CC | patient. The cancer that is diagnosed or treated is particularly |
| CC | prostate cancer. (I) and (II) can be used in vaccines. The antibodies or |
| CC | (I) can be used for monitoring the progression of cancer in a patient. |
| CC | (I) and (II) can also be used to improve diagnostic and therapeutic |
| CC | methods for prostate cancer. They can indicate the level of metastasis |
| CC | as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to |
| CC | AAH01318 represent polynucleotide and amino acid sequences used in the |
| CC | exemplification of the present invention. |
| XX | |
| XQ | Sequence 15 AA: |

| | | | | |
|-----------------------|-----------------|---------------|-----------|------------|
| Query Match | 39.3% | Score 11; | DB 22; | Length 15; |
| Best Local Similarity | 100.0% | Pred. No. 0 | 00034 | |
| Matches 11; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| OY | 16 | GTEARRHYDEG | 26 | |
| | | | | |
| | | | | |
| | | | | |
| db | 5 | gtearrhydeg | 15 | |

| | |
|----|--|
| XX | RESULT 11 |
| XX | AAG99053 |
| ID | AAG99053 standard; Peptide: 15 AA. |
| XX | |
| AC | AAG99053; |
| XX | |
| DT | 25-SEP-2001 (first entry) |
| XX | |
| DE | Human prostate-specific epitope mapping peptide #25. |
| XX | |
| KW | Human; prostate cancer; therapy; diagnosis; cat eye syndrome; |
| KW | chromosome 22q11.2; prostate-specific protein; chromosome 1; |
| KW | prostate specific antigen; PSA. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| PN | WO200134802-A2. |
| XX | |
| PD | 17-MAY-2001. |
| XX | |
| PF | 09-NOV-2000; 2000WO-US30904. |
| XX | |
| PR | 12-NOV-1999; 99US-0439313. |
| XX | |
| PR | 18-NOV-1999; 99US-0443686. |
| XX | |
| PA | (CORI-) CORIXA CORP. |
| XX | |
| PI | Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG; |
| PI | Kalos MD, Retler MM, Stolk JA, Day CH, Skelky YAW, Wang A; |
| DR | WPI; 2001-308785/32. |
| XX | |
| PT | Isolated polypeptide comprising at least an immunogenic portion of a |
| PT | prostate-specific protein, useful in the diagnosis and therapy of |
| XX | prostate cancer - |
| XX | |

PS Example 19; Page 302; 325pp; English.

XX

CC The present invention describes an isolated polypeptide (P1) comprising

CC at least an immunogenic portion of a prostate-specific protein, or its

CC variant. Also described are polynucleotides (N1) encoding (P1). (P1) and

CC (N1) have cytostatic activity and can be used in vaccine production.

CC The polypeptides, nucleic acids and antibodies from the present

CC invention are useful in the diagnosis and therapy of prostate cancer.

CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located

CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome

CC region. Prostate specific antigen (PSA) P501S was located on

CC chromosome 1. AAHH4671 to AAHH8143 and AAG39000 to AAG39077 represent

CC polynucleotide and polypeptide sequences used in the exemplification

CC of the present invention.

XX

SO Sequence 15 AA;

```

Query March 39.3% Score 11; DB 22; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 16 GTEARRHDEG 26
|||||111111111
db 5 gtearrhydeg 15

```

| | |
|----|---|
| XX | RESULT 12 |
| XX | AAU69814 |
| ID | AAU69814 standard; Peptide; 17 AA. |
| XX | |
| XX | AAU69814; |
| XX | |
| DT | 30-JAN-2002 (first entry) |
| XX | |
| DE | Human prostate protein P703P antigenic epitope #11. |
| XX | |
| KW | Human; prostate cancer; ss; cytostatic; immunostimulant; tumour; |
| KW | antigen; epitope. |
| XX | |
| OS | Homo sapiens. |
| OS | Synthetic. |
| PN | WO200173032-A2. |
| XX | |
| PD | 04-OCT-2001. |
| XX | |
| FF | 27-MAR-2001; 2001WO-US09919. |
| XX | |
| PR | 27-MAR-2000; 2000US-0536857. |
| PR | 09-MAY-2000; 2000US-0568100. |
| PR | 12-MAY-2000; 2000US-0570737. |
| PR | 13-JUN-2000; 2000US-0593793. |
| PR | 27-JUN-2000; 2000US-0605783. |
| PR | 10-AUG-2000; 2000US-0636213. |
| PR | 29-AUG-2000; 2000US-0651236. |
| PR | 06-SEP-2000; 2000US-0657279. |
| PR | 02-OCT-2000; 2000US-0679426. |
| PR | 10-OCT-2000; 2000US-0685166. |
| XX | |
| PA | (CORI-) CORIXA CORP. |
| XX | |
| PI | Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD; |
| PI | Fanger GR, Retter MW, Stolk JA, Day CH, Vedvyck TS, Carter D; |
| PI | Li SX, Wang A, Skelky YAM, Hepler WT, Henderson RA; |
| XX | |
| DR | WPI; 2001-639232/73. |
| XX | |
| PT | New human prostate-specific polypeptides and polynucleotides useful for |
| XX | the diagnosis and treatment of cancer, especially prostate cancer - |
| XX | Claim 2; Page 403; 579pp; English. |
| XX | |

CC The invention relates to isolated prostate-specific
CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of
CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC the antigen-presenting cells are useful for stimulating and/or expanding
CC T cells specific for a tumour protein, and for inhibiting the development
CC of cancer especially prostate cancer. Compositions comprising the
CC polynucleotide and/or polypeptide are useful for stimulating an immune
CC response, and for treating cancer. The oligonucleotide is useful for
CC detecting cancer. The present sequence is a prostate specific
CC peptide of the invention. The peptides either represent antigenic
CC epitopes or domains of prostate specific proteins.
CC
SQ Sequence 17 AA:

Query Match 21.4%; Score 6; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARR 21
|||||
Db 5 gtearr 10

RESULT 13
AAM01169
ID AAM01169 standard; Peptide: 17 AA.
XX
AC AAM01169:

DT 04-OCT-2001 (first entry)

DE Human prostate-specific peptide used in epitope mapping studies #26.

KW Human: prostate cancer; prostate-specific; diagnosis; vaccine;

KW cytosstatic; gene therapy; metastasis.

OS Homo sapiens.

PN WO200151633-A2.

PD 19-JUL-2001.

PF 16-JAN-2001; 2001WO-US01574.

PR 14-JAN-2000; 2000US-0483672.

PA (CORI-) CORIXA CORP.

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;

PI Wang A, Meagher MJ;

DR WPI; 2001-425873/45.

PT New polynucleotide encoding a prostate-specific protein, for

PT diagnosing, monitoring and treating prostate cancer in a patient and

PT for use in vaccines -

PS Claim 2; Page 402-403; 543pp; English.

CC The present invention describes polynucleotide sequences (I) which encode

CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,

CC and can be used in vaccine production and gene therapy. (II), (III),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.

CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC AAM01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
CC
SQ Sequence 17 AA:

Query Match 21.4%; Score 6; DB 22; Length 17;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARR 21
|||||
Db 5 gtearr 10

RESULT 14
AAG99054
ID AAG99054 standard; Peptide: 17 AA.
XX
AC AAG99054:

DT 25-SEP-2001 (first entry)

DE Human prostate-specific epitope mapping peptide #26.

KW Human: prostate cancer; therapy; diagnosis; cat eye syndrome;

KW Chromosome 22q11.2; prostate-specific protein; chromosome 1;

KW prostate specific antigen; PSA.

OS Homo sapiens.

PN WO200134802-A2.

PD 17-MAY-2001.

PF 09-NOV-2000; 2000WO-US30904.

PR 12-NOV-1999; 99US-0439313.

PR 18-NOV-1999; 99US-0443686.

PA (CORI-) CORIXA CORP.

PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;

PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;

PI WPI; 2001-308785/32.

PT Isolated polypeptide comprising at least an immunogenic portion of a

PT prostate-specific protein, useful in the diagnosis and therapy of

PT prostate cancer -

PS Claim 3; Page 302; 325pp; English.

CC The present invention describes an isolated polypeptide (PI) comprising

CC at least an immunogenic portion of a prostate-specific protein, or its

CC variant. Also described are polynucleotides (NI) encoding (PI) and

CC (NI) have cytostatic activity and can be used in vaccine production.
CC The polypeptides, nucleic acids and antibodies from the present
CC invention are useful in the diagnosis and therapy of prostate cancer.
CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC in a genomic region on chromosome 22q11.2 known as the Cat Eye Syndrome
CC region. Prostate specific antigen (PSA) P501S was located on
CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC polynucleotide and polypeptide sequences used in the exemplification
CC of the present invention.
CC
SQ Sequence 17 AA:

Query Match 21.4%; Score 6; DB 22; Length 17;

Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEARR 21
|||||
Db 5 gtearr 10

RESULT 15

AAE01151
ID AAE01151 standard; Protein: 8 AA.

AC AAE01151;

DT 17-JUL-2001 (first entry)

DE Peptide #1 of peppermint GPP synthase small subunit from clone pMp13.18.

KW Peppermint; cytostatic; geranyl diphosphate synthase large subunit; GPP;

KM prenyltransferase; terpenoid family; dimethylallyl diphosphate; DMAPP;

KW isopentenyl diphosphate; IPP; acyclic precursor; monoterpene synthase;

XX limonene synthase; therapy; cancer; gene therapy; clone pMp13.18.

OS Mentha piperita.

PN WO200129188-A1.

PD 26-APR-2001.

PF 16-OCT-2000; 2000WO-US28690.

PR 18-OCT-1999; 99US-0420211.

PA (UNIV) UNIV WASHINGTON STATE RES FOUND.

PI Croteau RB, Burke CC, Wildung MR;

DR WPI; 2001-300331/31.

PT New isolated recombinant geranyl diphosphate synthase large subunit

XX protein, useful for treating cancer

PS Example 1; Page 36; 68pp; English.

CC The present sequence is a peptide of peppermint geranyl diphosphate

CC (GPP) synthase small subunit from clone pMp13.18. GPP synthase is one

CC of a family of enzymes called prenyltransferases that catalyze C5

CC elongation reactions to form the linear (acyclic) precursors of the

CC various terpenoid families. GPP synthase catalyzes the condensation of

CC dimethylallyl diphosphate (DMAPP) and isopentenyl diphosphate (IPP) to

CC form GPP which is the intermediate, C10 acyclic precursor of the

CC monoterpenes. GPP synthase along with a monoterpene synthase protein,

CC preferably limonene synthase, which is capable of converting geranyl

CC diphosphate to a monoterpene having anti-cancer properties, is useful for

CC treating cancer in a mammalian host by introducing GPP synthase and the

CC monoterpene synthase protein into a cancerous cell. GPP synthase large

CC subunit DNA along with DNAs encoding GPP synthase small subunit and

CC monoterpene synthase protein is useful for treating cancer. GPP synthase

CC DNA is used in gene therapy.

Search completed: June 26, 2002, 13:51:47
Job time: 165 sec

XX Sequence 8 AA;

Query Match 17.9%; Score 5; DB 22; Length 8;

Best Local Similarity 100.0%; Pred. No. 6.4e+05; Indels 0; Gaps 0;

Matches 5; Conservative 0; Mismatches 0;

OY 6 GLYOG 10
|||||
Db 2 glyog 6

Thu Jun 27 09:46:45 2002

us-09-838-785-25.oli.rag

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:49:32 ; Search time 14.89 seconds
(without alignments)
45.931 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28

Sequence: 1 DFVGEGLYGVPRAGCTEARHRYDEGVR 28

Scoring table:

GAPOP 60.0 , GAPEXT 60.0

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 140943

Minimum DB seq length: 0
Maximum DB seq length: 28

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCtUS.COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Length | DB ID | Description |
|------------|-------|---------|--------|-------|--------------------|
| 1 | 13 | 46.4 | 25 | 4 | US-09-439-313-520 |
| 2 | 13 | 46.4 | 27 | 4 | US-09-071-710-39 |
| 3 | 13 | 46.4 | 27 | 4 | US-09-525-397-39 |
| 4 | 13 | 46.4 | 27 | 4 | US-09-439-313-566 |
| 5 | 11 | 39.3 | 15 | 4 | US-09-439-313-518 |
| 6 | 6 | 21.4 | 17 | 4 | US-09-439-313-519 |
| 7 | 5 | 17.9 | 8 | 4 | US-09-420-211-12 |
| 8 | 5 | 17.9 | 9 | 2 | US-08-951-924A-3 |
| 9 | 5 | 17.9 | 11 | 3 | US-08-593-500-42 |
| 10 | 5 | 17.9 | 11 | 3 | US-08-195-006-42 |
| 11 | 5 | 17.9 | 11 | 5 | PCT-US94-07644A-42 |
| 12 | 5 | 17.9 | 14 | 1 | US-08-277-299-1 |
| 13 | 5 | 17.9 | 14 | 1 | US-08-880-671-1 |
| 14 | 5 | 17.9 | 14 | 4 | US-08-853-910-5 |
| 15 | 5 | 17.9 | 15 | 1 | US-07-969-336-1 |
| 16 | 5 | 17.9 | 15 | 2 | US-08-815-953-1 |
| 17 | 5 | 17.9 | 17 | 2 | US-08-472-659-4 |
| 18 | 5 | 17.9 | 17 | 2 | US-08-474-661-4 |
| 19 | 5 | 17.9 | 17 | 2 | US-08-611-977-4 |
| 20 | 5 | 17.9 | 18 | 1 | US-07-920-281C-27 |
| 21 | 5 | 17.9 | 18 | 4 | US-08-466-277-27 |
| 22 | 5 | 17.9 | 19 | 4 | US-08-975-040-8 |
| 23 | 5 | 17.9 | 26 | 2 | US-08-563-892A-1 |
| 24 | 5 | 17.9 | 26 | 2 | US-08-563-892A-20 |
| 25 | 5 | 17.9 | 26 | 2 | US-08-563-892A-21 |
| 26 | 5 | 17.9 | 26 | 2 | US-08-563-892A-22 |
| 27 | 4 | 14.3 | 4 | 2 | US-08-248-839C-21 |

| | | | | | | |
|----|---|------|---|---|--------------------|--------------------|
| 28 | 4 | 14.3 | 4 | 2 | US-08-623-833B-24 | Sequence 24, Appl |
| 29 | 4 | 14.3 | 5 | 6 | 5166318-10 | Patent No. 5166318 |
| 30 | 4 | 14.3 | 6 | 1 | US-08-211-747-4 | Sequence 4, Appl |
| 31 | 4 | 14.3 | 6 | 3 | US-09-146-675-7 | Sequence 7, Appl |
| 32 | 4 | 14.3 | 8 | 1 | US-08-178-570-54 | Sequence 54, Appl |
| 33 | 4 | 14.3 | 8 | 1 | US-08-071-357-14 | Sequence 14, Appl |
| 34 | 4 | 14.3 | 8 | 1 | US-08-634-060-40 | Sequence 40, Appl |
| 35 | 4 | 14.3 | 8 | 1 | US-08-211-747-1 | Sequence 1, Appl |
| 36 | 4 | 14.3 | 8 | 2 | US-08-807-030-29 | Sequence 29, Appl |
| 37 | 4 | 14.3 | 8 | 2 | US-08-807-030-53 | Sequence 53, Appl |
| 38 | 4 | 14.3 | 8 | 3 | US-08-369-643-54 | Sequence 54, Appl |
| 39 | 4 | 14.3 | 8 | 5 | PCT-US95-00147-54 | Sequence 54, Appl |
| 40 | 4 | 14.3 | 9 | 1 | US-08-467-083-1 | Sequence 1, Appl |
| 41 | 4 | 14.3 | 9 | 1 | US-08-211-747-8 | Sequence 8, Appl |
| 42 | 4 | 14.3 | 9 | 1 | US-08-414-417B-1 | Sequence 1, Appl |
| 43 | 4 | 14.3 | 9 | 2 | US-08-486-348A-1 | Sequence 1, Appl |
| 44 | 4 | 14.3 | 9 | 2 | US-08-468-545B-1 | Sequence 1, Appl |
| 45 | 4 | 14.3 | 9 | 3 | US-08-159-339A-741 | Sequence 741, App |

ALIGNMENTS

```
RESULT 1
US-09-439-313-520
; Sequence 520, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yugu
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.427C9
; CURRENT APPLICATION NUMBER: US/09/439,313
; CURRENT FILING DATE: 1999-11-12
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 520
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Made in a lab
; US-09-439-313-520

Query Match      46.4% Score 13: DB 4: Length 25:
Best Local Similarity 100.0% Pred. No. 1.7e-06:
Matches 13: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY      3 VEGGLYGVPRAE 15
DB      1 VEGGLYGVPRAE 13

RESULT 2
US-09-071-710-39
; Sequence 39, Application US/09071710
; Patent No. 6130043
; GENERAL INFORMATION:
; APPLICANT: BILLING-MEDEL, PATRICIA
; APPLICANT: COHEN, MAURICE
; APPLICANT: COLPITTS, TRACEY L.
; APPLICANT: FRIEDMAN, PAULA N.
```

APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850,713
FILING DATE: 02-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6130043e
US-09-071-710-39

Query Match 46.4%; Score 13; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEARRHYDEGVR 28
|||||

DB 12 GTEARRHYDEGVR 24

RESULT 3
US-09-525-397-39
Sequence 39, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.

APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525,397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-39

Query Match 46.4%; Score 13; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEARRHYDEGVR 28
|||||

DB 12 GTEARRHYDEGVR 24

RESULT 4
US-09-439-313-566
Sequence 566, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqul
APPLICANT: Reed, Steven G.
APPLICANT: KALOS, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FASTSEQ for Windows Version 3.0

SEQ ID NO 566
LENGTH: 27
TYPE: PRT
ORGANISM: Homo sapiens
US-09-439-313-566

Query Match
Best Local Similarity 100.0%; Score 13; DB 4; Length 27;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VEGGLYQGVPAE 15
DB 1 VEGGLYQGVPAE 13

RESULT 5
US-09-439-313-518
Sequence 518, Application US/09439313
Patent No. 6329505

GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqiu
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 518
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-439-313-518

Query Match
Best Local Similarity 100.0%; Score 11; DB 4; Length 15;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARRHYDEG 26
DB 5 GTEARRHYDEG 15

RESULT 6
US-09-439-313-519
Sequence 519, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqiu
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Craig

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 519
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-439-313-519

Query Match
Best Local Similarity 100.0%; Score 6; DB 4; Length 17;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARR 21
DB 5 GTEARR 10

RESULT 7
US-09-420-211-12
Sequence 12, Application US/09420211B
Patent No. 6303330
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Burke, Charles C
APPLICANT: Wildung, Mark R
TITLE OF INVENTION: GERANYL DIPHOSPHATE SYNTHASE LARGE SUBUNIT, AND METHODS
FILE REFERENCE: WSUR14294
CURRENT APPLICATION NUMBER: US/09/420,211B
CURRENT FILING DATE: 1999-10-18
EARLIER APPLICATION NUMBER: PCT/US98/21772
EARLIER FILING DATE: 1998-10-15
EARLIER APPLICATION NUMBER: 08/951,924
EARLIER FILING DATE: 1997-10-16
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 8
TYPE: PRT
ORGANISM: Mentha piperita
US-09-420-211-12

Query Match
Best Local Similarity 100.0%; Score 5; DB 4; Length 8;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GLYOG 10
DB 2 GLYOG 6

RESULT 8
US-08-951-924A-3
Sequence 3, Application US/08951924A
Patent No. 5876964
GENERAL INFORMATION:
APPLICANT: Croteau, Rodney B
APPLICANT: Wildung, Mark R
APPLICANT: Burke, Charles
APPLICANT: Gershenzon, Jonathan
TITLE OF INVENTION: GERANYL DIPHOSPHATE SYNTHASE FROM MINT
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:

ADDRESSEE: CHRISTENSEN, O'CONNOR, JOHNSON AND KINDNESS
STREET: 1420 FIFTH AVENUE
CITY: SEATTLE
STATE: WASHINGTON
COUNTRY: USA
ZIP: WA 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/951,924A
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Shelton, Dennis K
REGISTRATION NUMBER: 26,997
REFERENCE/DOCKET NUMBER: WSUR 1 11217
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206 695 1718
TELEFAX: 206 224 0779
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Mencha spicata
US-08-951-924A-3

Query Match 17.9%; Score 5; DB 2; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.7e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 6 GLYOG 10
11111
DB 2 GLYOG 6

RESULT 9
US-08-592-500-42
Sequence 42, Application US/08592500
Patent No. 6005089
GENERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
APPLICANT: Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/592,500
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11
OTHER INFORMATION: /note="Amino acid sequence of the
OTHER INFORMATION: human plasma factor XIII (F XIII) thrombin cleavage
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: Region
LOCATION: 6..8
OTHER INFORMATION: /note="Amino acid residues
OTHER INFORMATION: identical to GPV."
US-08-592-500-42

Query Match 17.9%; Score 5; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0;

OY 9 OGVR 13
11111
DB 3 OGVR 7

RESULT 10
US-08-195-006-42
Sequence 42, Application US/08195006
Patent No. 6083688
GENERAL INFORMATION:
APPLICANT: Lanza, Francois
APPLICANT: Phillips, David R.
APPLICANT: Cazenave, Jean-Pierre
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourile and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/195,006
FILING DATE: 10-FEB-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/089,455
FILING DATE: 09-JUL-1993
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 12418-28
TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11 /note="Amino acid sequence of the
OTHER INFORMATION: human plasma factor XIII (FXIII) thrombin cleavage
OTHER INFORMATION: site."
FEATURE:
NAME/KEY: Region
LOCATION: 6..8
OTHER INFORMATION: /note="Amino acid residues
OTHER INFORMATION: identical to GPV."
US-08-195-006-42

Query Match 17.9%; Score 5; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 OGVR 13
|||||
DB 3 OGVR 7

RESULT 11
PCT-US94-07644A-42
Sequence 42, Application PC/TUS9407644A
GENERAL INFORMATION:
APPLICANT: COR Therapeutics, Inc.
TITLE OF INVENTION: Platelet Glycoprotein V Gene and Uses
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/07644A
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 012418-003000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 326-2400
TELEFAX: (415) 326-2422
INFORMATION FOR SEQ ID NO: 42:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..11 /note="Amino acid sequence of the
OTHER INFORMATION: human plasma factor XIII (FXIII) thrombin cleavage
OTHER INFORMATION: human plasma factor XIII (FXIII) thrombin cleavage

OTHER INFORMATION: site."
FEATURE:
NAME/KEY: Region
LOCATION: 6..8
OTHER INFORMATION: /note="Amino acid residues
OTHER INFORMATION: identical to GPV."
PCT-US94-07644A-42

Query Match 17.9%; Score 5; DB 5; Length 11;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 OGVR 13
|||||
DB 3 OGVR 7

RESULT 12
US-08-277-299-1
Sequence 1, Application US/08277299
GENERAL INFORMATION:
APPLICANT: Ochoa, Augusto C.
APPLICANT: Young, Howard A.
APPLICANT: Longo, Dan L.
APPLICANT: Ghosh, Paritosh
APPLICANT: Robb, Richard J.
APPLICANT: Neville, Mary E.
TITLE OF INVENTION: METHODS OF IDENTIFYING PATIENTS HAVING
TITLE OF INVENTION: ALTERED IMMUNE STATUS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/277,299
FILING DATE: 22-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SAXE, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 40403/132/ONCO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-277-299-1

Query Match 17.9%; Score 5; DB 1; Length 14;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GLYOG 10
|||||
DB 9 GLYOG 13

RESULT 13
US-08-880-671-1
Sequence 1, Application US/08880671
Patent No. 5965366
GENERAL INFORMATION:
APPLICANT: Ochoa, Augusto C.
APPLICANT: Young, Howard A.
APPLICANT: Longo, Dan L.
APPLICANT: Ghosh, Paritosh
APPLICANT: Robb, Richard J.
APPLICANT: Neville, Mary E.
TITLE OF INVENTION: METHODS OF IDENTIFYING PATIENTS HAVING
TITLE OF INVENTION: ALTERED IMMUNE STATUS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/880,671
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/277,299
FILING DATE: 22-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: Saxe, Bernhard D.
REGISTRATION NUMBER: 28,665
REFERENCE/DOCKET NUMBER: 40403/132/ONCO
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-880-671-1

Query Match 17.9%; Score 5; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GLYOG 10
|||||
DB 9 GLYOG 13

RESULT 14
US-08-853-910-5
Sequence 5, Application US/08853910
Patent No. 6291247
GENERAL INFORMATION:
APPLICANT: Ripollé, Richard J.
APPLICANT: Ross, Gregory M.
APPLICANT: Dory, Magdalena I.
APPLICANT: Weaver, Donald F.
APPLICANT: Shamovsky, Igor L.
TITLE OF INVENTION: Factors and Methods for Reducing Biological
TITLE OF INVENTION: Activity of a Multimeric Protein and Methods of Screening
TITLE OF INVENTION: Therefor
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Dowell & Dowell
STREET: 2001 Jefferson Davis Highway
CITY: Arlington
STATE: Virginia
COUNTRY: United States of America
ZIP: 22202-3697
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/853,910
FILING DATE: 09-MAY-1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 08/241462
FILING DATE: 11-MAY-1994
APPLICATION NUMBER: USSN 60/010 328
FILING DATE: 09-NOV-1995
APPLICATION NUMBER: USSN 08/745608
FILING DATE: 08-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: RALPH A. DOWELL
REGISTRATION NUMBER: 26868
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 415-2555
TELEFAX: (703) 415-2559
INFORMATION FOR SEO ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 14 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-853-910-5

Query Match 17.9%; Score 5; DB 4; Length 14;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 VPRAE 15
|||||
DB 5 VPRAE 9

RESULT 15
US-07-969-336-1
Sequence 1, Application US/07969336
Patent No. 5681729
GENERAL INFORMATION:
APPLICANT: Kudo, Toshio
APPLICANT: Hiyanuta, Toshiaki
APPLICANT: Tanedani, Toshiyuki
APPLICANT: Kadota, Akihiko
TITLE OF INVENTION: Methods for Producing Human Lymphocytes
TITLE OF INVENTION: and Human Monoclonal Antibodies, and Human Monoclonal
TITLE OF INVENTION: Antibodies Produced Thereby
NUMBER OF SEQUENCES: 1
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Birch, Stewart, Kolasch and Birch
STREET: 301 No. 5681729th Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046-3487
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/969,336

FILING DATE: 19921030
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Raymond C.
 REGISTRATION NUMBER: 21,066
 REFERENCE/DOCKET NUMBER: 760-159P
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 241-1300
 TELEFAX: (703) 241-2848
 TELEX: 248345
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 15 amino acids
 TYPE: AMINO ACID
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 HYPOTHETICAL: NO
 FRAGMENT TYPE: N-terminal
 US-07-969-336-1

Query Match 17.98; Score 5; DB 1; Length 15;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 VEGGL 7
 Db 11 VEGGL 15

Search completed: June 26, 2002, 13:52:14
 Job time: 162 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:50:07 : Search time 20.27 Seconds
(without alignments)
132.733 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28

Sequence: 1 DFVGEGLYGVPRAGEFRRHRYDEGVR 28

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5685

Minimum DB seq length: 0
Maximum DB seq length: 28

Post-processing: Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------------|
| 1 | 4 | 14.3 | 8 | 2 | B45800 serum albumin - do |
| 2 | 4 | 14.3 | 11 | 2 | D60691 phycobillosome 9K I |
| 3 | 4 | 14.3 | 12 | 2 | S47393 T-cell antigen rec |
| 4 | 4 | 14.3 | 13 | 2 | G22565 R-phycoerythrin ga |
| 5 | 4 | 14.3 | 14 | 2 | PH1677 Ig heavy chain V r |
| 6 | 4 | 14.3 | 16 | 2 | A36285 site-specific meth |
| 7 | 4 | 14.3 | 16 | 2 | PH0754 T-cell receptor be |
| 8 | 4 | 14.3 | 17 | 2 | A60317 glucagon-like pept |
| 9 | 4 | 14.3 | 18 | 2 | A24749 neuropeptide A - b |
| 10 | 4 | 14.3 | 20 | 2 | A60525 lysozyme (EC 3.2.1 |
| 11 | 4 | 14.3 | 20 | 2 | A40451 dormancy-related p |
| 12 | 4 | 14.3 | 20 | 2 | I58192 glial fibrillary a |
| 13 | 4 | 14.3 | 21 | 2 | PH1688 Ig heavy chain V r |
| 14 | 4 | 14.3 | 21 | 2 | PH1690 Ig heavy chain V r |
| 15 | 4 | 14.3 | 22 | 2 | PH1678 Ig heavy chain V r |
| 16 | 4 | 14.3 | 22 | 2 | PH1679 Ig heavy chain V r |
| 17 | 4 | 14.3 | 22 | 2 | PH1680 Ig heavy chain V r |
| 18 | 4 | 14.3 | 22 | 2 | A37335 68k surface antige |
| 19 | 4 | 14.3 | 22 | 2 | S03518 T-cell receptor ga |
| 20 | 4 | 14.3 | 22 | 2 | PT0052 translatin initia |
| 21 | 4 | 14.3 | 23 | 2 | PH1681 Ig heavy chain V r |
| 22 | 4 | 14.3 | 23 | 2 | PH1682 Ig heavy chain V r |
| 23 | 4 | 14.3 | 23 | 2 | PH1689 Ig heavy chain V r |
| 24 | 4 | 14.3 | 23 | 2 | PH1691 Ig heavy chain V r |
| 25 | 4 | 14.3 | 23 | 2 | PH1692 Ig heavy chain V r |
| 26 | 4 | 14.3 | 23 | 2 | PH1693 Ig heavy chain V r |
| 27 | 4 | 14.3 | 23 | 2 | PH1694 Ig heavy chain V r |
| 28 | 4 | 14.3 | 24 | 2 | PH1683 Ig heavy chain V r |
| 29 | 4 | 14.3 | 24 | 2 | PH1684 Ig heavy chain V r |

| | | | | | |
|----|---|------|----|---|---------------------------|
| 30 | 4 | 14.3 | 24 | 2 | PH1685 Ig heavy chain V r |
| 31 | 4 | 14.3 | 25 | 2 | PH1686 Ig heavy chain V r |
| 32 | 4 | 14.3 | 25 | 2 | PH1700 Ig heavy chain V r |
| 33 | 4 | 14.3 | 25 | 2 | PH1701 Ig heavy chain V r |
| 34 | 4 | 14.3 | 26 | 2 | PH1687 Ig heavy chain V r |
| 35 | 4 | 14.3 | 26 | 2 | PH1703 Ig heavy chain V r |
| 36 | 4 | 14.3 | 26 | 2 | PH1704 Ig heavy chain V r |
| 37 | 4 | 14.3 | 26 | 2 | S27117 macrophage migrati |
| 38 | 4 | 14.3 | 27 | 4 | JE0015 probable 2.8k prot |
| 39 | 4 | 14.3 | 27 | 4 | JS0104 probable 2.8k prot |
| 40 | 3 | 10.7 | 5 | 2 | B22565 R-phycoerythrin a1 |
| 41 | 3 | 10.7 | 6 | 2 | PT0512 T-cell receptor be |
| 42 | 3 | 10.7 | 6 | 2 | PT0648 T-cell receptor be |
| 43 | 3 | 10.7 | 7 | 2 | PH1408 Ig heavy chain V r |
| 44 | 3 | 10.7 | 7 | 2 | A11483 aspartate transami |
| 45 | 3 | 10.7 | 7 | 2 | PT0529 T-cell receptor be |

ALIGNMENTS

RESULT 1
B45800 serum albumin - dog (fragment)
C:Species: Canis lupus familiaris (dog)
C>Date: 03-Jun-1993 #sequence-revision 03-Jun-1993 #text-change 31-Dec-1993
C:Accession: B45800
R:Carraway, R.E.; Cochran, D.E.; Boucher, W.; Mitra, S.P.
J. Immunol. 143, 1680-1684, 1989
A:Title: Structures of histamine-releasing peptides formed by the action of acid prot
A:Reference number: A45800; MUID:89341406
A:Accession: B45800
A>Status: preliminary
A:Molecule type: protein
A:Residues: 1-8 <CAR>

Query Match 14.3%; Score 4; DB 2; Length 8;
Best Local Similarity 100.0%; Pred. No. 2.8e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 ARRH 22
DB 2 ARRH 5

RESULT 2
D60691 phycobillosome 9K linker protein - Synechococcus sp. (PCC 7002) (fragment)
C:Species: Synechococcus sp.
C>Date: 14-May-1993 #sequence-revision 14-May-1993 #text-change 07-May-1999
C:Accession: D60691
R:Bryant, D.A.; de Lormier, R.; Guglielmi, G.; Stevens Jr., S.E.
Arch. Microbiol. 153, 550-560, 1990
A:Title: Structural and compositional analyses of the phycobillosomes of Synechococcus
S:Reference number: A60691; MUID:90314662
A:Accession: D60691
A:Molecule type: protein
A:Residues: 1-11 <RRY>
A:Comment: This protein, one of the eleven components detected in this species of the
C:Keywords: photosystem II

Query Match 14.3%; Score 4; DB 2; Length 11;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
DB 8 GTEA 11

RESULT 3
S47393
T-cell antigen receptor VJ junction beta chain - human
C:Species: Homo sapiens (man)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
C:Accession: S47393
R:Rehner, P.J.
submitted to the EMBL Data Library, August 1994
A:Description: Human HLA-A0201 restricted recognition of Influenza A is dominated by T
A:Reference number: S47355
A:Accession: S47393
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-12 <LEH>
A:Cross-references: EMBL:Z53670; NID:G527445; PIDN:CAA84739.1; PID:G527446
C:Keywords: T-cell receptor

Query Match 14.3%; Score 4; DB 2; Length 12;
Best Local Similarity 100.0%; Pred. No. 5.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
|||||
Db 7 GTEA 10

RESULT 4
G22565
R-phycoerythrin gamma-B chain - red alga (Gastroclonium coulteri) (fragment)
C:Species: Gastroclonium coulteri
C:Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C:Accession: G22565
R:Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A:Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A:Reference number: A22565; MUID:85182601
A:Accession: G22565
A:Molecule type: Protein
A:Residues: 1-13 <KLO>

Query Match 14.3%; Score 4; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 AEGT 17
|||||
Db 2 AEGT 5

RESULT 5
PH1677
Ig heavy chain V region (clone NP-6-6) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: PH1677
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vivo.
A:Reference number: PH1675; MUID:93301607
A:Accession: PH1677
A:Molecule type: mRNA
A:Residues: 1-14 <MCH>
A:Experimental source: B cell
C:Superfamily: Immunoglobulin V region; Immunoglobulin homology
C:Keywords: heterotetramer; Immunoglobulin

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 TEAR 20
|||||
Db 11 TEAR 14

RESULT 6
A36285
site-specific methyltransferase (EC 2.1.1.-) - Escherichia coli (fragment)
C:Species: Escherichia coli
C:Date: 18-Jan-1991 #sequence_revision 18-Jan-1991 #text_change 21-Mar-1996
C:Accession: A36285
R:Reich, N.O.; Everett, E.A.
J. Biol. Chem. 265, 8929-8934, 1990
A:Title: Identification of peptides involved in S-adenosylmethionine binding in the E
A:Reference number: A36285; MUID:90256827
A:Accession: A36285
A:Status: Preliminary
A:Molecule type: protein
A:Residues: 1-16 <REI>
C:Keywords: methyltransferase

Query Match 14.3%; Score 4; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
|||||
Db 1 GTEA 4

RESULT 7
PH0754
T-cell receptor beta chain (F1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 05-Nov-1999
C:Accession: PH0754
R:Casanova, J.L.; Romero, P.; Widmann, C.; Kourilsky, P.; Maryanski, J.L.
J. Exp. Med. 174, 1371-1383, 1991
A:Title: T cell receptor genes in a series of class I major histocompatibility complex
allelic exclusion and antigen-specific repertoire.
A:Reference number: PH0746; MUID:92078846
A:Accession: PH0754
A:Molecule type: mRNA
A:Residues: 1-16 <CAS>
A:Cross-references: EMBL:X60847; NID:G50935; PIDN:CAA43238.1; PID:G50936
A:Experimental source: T lymphocyte
C:Keywords: T-cell receptor

Query Match 14.3%; Score 4; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 6.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 24 DEGV 27
|||||
Db 5 DEGV 8

RESULT 8
A60317
glucagon-like peptide I - marbled electric ray (fragment)
C:Species: Torpedo marmorata (marbled electric ray)
C:Date: 28-Oct-1992 #sequence_revision 28-Oct-1992 #text_change 21-Nov-1997
C:Accession: A60317
R:Conlon, J.M.; Hansen, H.F.; Schwartz, T.W.
Regul. Pept. 13, 94, 1986
A:Title: A truncated glucagon-like peptide I from torpedo pancreas.
A:Reference number: A60317
A:Accession: A60317
A:Molecule type: Protein
A:Residues: 1-17 <CON>
C:Superfamily: glucagon

C:Keywords: duplication; pancreas

Query Match 14.3%; Score 4; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 AEGT 17
||||
DB 2 AEGT 5

RESULT 9
A24749

neuropeptide A - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 28-Jul-1987 #sequence_revision 28-Jul-1987 #text_change 18-Aug-2000
C:Accession: A24749
R:Yang, H.Y.T.; Pratta, W.; Majane, E.A.; Costa, E.
Proc. Natl. Acad. Sci. U.S.A. 82, 7757-7761, 1985
A:Title: Isolation, sequencing, synthesis, and pharmacological characterization of two h
A:Reference number: A94074; MUID:86067985
A:Accession: A24749
A:Molecule type: protein
A:Residues: 1-18 <YAN>
C:Comment: The source of this peptide was brain.
C:Superfamily: unassigned animal peptides
C:Keywords: neuropeptide

Query Match 14.3%; Score 4; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 7.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GEG 7
||||
DB 2 GEG 5

RESULT 10
A60525

lysozyme (EC 3.2.1.17) - cat (fragment)
C:Species: Felis silvestris catus (domestic cat)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 20-Jun-2000
C:Accession: A60525
R:Halliday, J.A.; Bell, K.; McKenzie, H.A.; Shaw, D.C.
Comp. Biochem. Physiol. B 95, 773-779, 1990
A:Title: Feline whey proteins: identification, isolation and initial characterization of
A:Reference number: A60525; MUID:90263403
A:Accession: A60525
A:Molecule type: protein
A:Residues: 1-20 <HAL>
A:Experimental source: milk
C:Superfamily: lysozyme c
C:Keywords: bacteriolytic enzyme; glycosidase; hydrolase; polysaccharide degradation

Query Match 14.3%; Score 4; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RAEG 16
||||
DB 13 RAEG 16

RESULT 11
A40451

dormancy-related protein Pin 1 I - sugar pine (fragment)
C:Species: Pinus lambertiana (sugar pine)
C:Date: 06-Dec-1991 #sequence_revision 06-Dec-1991 #text_change 18-Jun-1993
C:Accession: A40451
R:Ekrumodoullah, A.K.M.

submitted to the Protein Sequence Database, November 1991
A:Reference number: A40451

A:Accession: A40451
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <EKR>

Query Match 14.3%; Score 4; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 EARR 21
||||
DB 14 EARR 17

RESULT 12
I58192

glial fibrillary acidic protein GAP - rat (fragment)
C:Species: Rattus sp. (rat)
C:Date: 27-Feb-1997 #sequence_revision 27-Feb-1997 #text_change 27-Feb-1997
C:Accession: I58192
R:Laping, N.J.; Morgan, T.E.; Nichols, N.R.; Rozovsky, I.; Young-Chan, C.S.; Zarow, C.
Neuroscience 58, 563-572, 1994
A:Title: Transforming growth factor-beta 1 induces neuronal and astrocyte genes: tubu
A:Reference number: I58192; MUID:94224314
A:Accession: I58192
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-20 <RES>
A:Cross-references: GB:S70248; NID:9546523
A:Experimental source: strain Fisher
C:Genetics:
A:introns: 11/2

Query Match 14.3%; Score 4; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 8.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 EARR 21
||||
DB 8 EARR 11

RESULT 13
P1688

lg heavy chain V region (clone NP-7-1) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999
C:Accession: P1688
R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.
J. Exp. Med. 178, 295-307, 1993
A:Title: Antigen-driven B cell differentiation in vlvo.
A:Reference number: P1675; MUID:93301607
A:Accession: P1688
A:Molecule type: mRNA
A:Residues: 1-21 <MCH>
A:Experimental source: B cell
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin

Query Match 14.3%; Score 4; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 8.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 TEAR 20
||||
DB 11 TEAR 14

RESULT 14

PH1690

Ig heavy chain V region (clone NP-7-3) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C:Accession: PH1690

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PH1675; MUID:93301607

A:Accession: PH1690

A:Molecule type: mRNA

A:Residues: 1-21 <MCH>

A:Experimental source: B cell

A:Note: the authors translated the codon AAT for residue 9 as Lys

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match

14.3%; Score 4; DB 2; Length 21;

Best Local Similarity 100.0%; Pred. No. 8.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TEAR 20

||||

Db 11 TEAR 14

RESULT 15

PH1678

Ig heavy chain V region (clone NP-6-9) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 24-Feb-1994 #sequence_revision 24-Feb-1994 #text_change 17-Mar-1999

C:Accession: PH1678

R:McHeyzer-Williams, M.G.; McLean, M.J.; Lalor, P.A.; Nossal, G.J.V.

J. Exp. Med. 178, 295-307, 1993

A:Title: Antigen-driven B cell differentiation in vivo.

A:Reference number: PH1675; MUID:93301607

A:Accession: PH1678

A:Molecule type: mRNA

A:Residues: 1-22 <MCH>

A:Experimental source: B cell

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

Query Match

14.3%; Score 4; DB 2; Length 22;

Best Local Similarity 100.0%; Pred. No. 8.9e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 TEAR 20

||||

Db 11 TEAR 14

Search completed: June 26, 2002, 13:52:47
Job time: 160 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:51:08 ; Search time 9.93 Seconds

(without alignments)
109.179 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28
Sequence: 1 DFVGEGLYGVPRAGTEARRHYDEGVR 28

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1673

Minimum DB seq length: 0

Maximum DB seq length: 28

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 4 | 14.3 | 9 | 1 | NEUX_HUMAN |
| 2 | 4 | 14.3 | 18 | 1 | NPA_BOVIN |
| 3 | 4 | 14.3 | 20 | 1 | FIIB_BOVIN |
| 4 | 4 | 14.3 | 20 | 1 | LYC_FELCA |
| 5 | 4 | 14.3 | 20 | 1 | MYF_PIG |
| 6 | 4 | 14.3 | 21 | 1 | FIIB_BISBO |
| 7 | 4 | 14.3 | 21 | 1 | FIIB_BUBBU |
| 8 | 4 | 14.3 | 21 | 1 | FIIB_SYNCA |
| 9 | 4 | 14.3 | 21 | 1 | OMP1_ACPPL |
| 10 | 4 | 14.3 | 22 | 1 | IF2G_PIG |
| 11 | 4 | 14.3 | 28 | 1 | CH60_MYCSM |
| 12 | 4 | 10.7 | 5 | 1 | TPIS_CANFA |
| 13 | 3 | 10.7 | 8 | 1 | UPAI_HUMAN |
| 14 | 3 | 10.7 | 8 | 1 | VGIG_HSV2B |
| 15 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 16 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 17 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 18 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 19 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 20 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 21 | 3 | 10.7 | 9 | 1 | FIIB_PAPAN |
| 22 | 3 | 10.7 | 10 | 1 | TKL2_LOCM1 |
| 23 | 3 | 10.7 | 10 | 1 | TKL3_LOCM1 |
| 24 | 3 | 10.7 | 10 | 1 | TKL4_LOCM1 |
| 25 | 3 | 10.7 | 10 | 1 | TKNB_CHICK |
| 26 | 3 | 10.7 | 10 | 1 | TKNB_ONCMY |
| 27 | 3 | 10.7 | 10 | 1 | TKNB_PIG |
| 28 | 3 | 10.7 | 10 | 1 | TKNB_PIG |
| 29 | 3 | 10.7 | 10 | 1 | TKNB_PIG |
| 30 | 3 | 10.7 | 10 | 1 | TKNB_PIG |
| 31 | 3 | 10.7 | 10 | 1 | TKNB_PIG |
| 32 | 3 | 10.7 | 10 | 1 | TKNB_PIG |
| 33 | 3 | 10.7 | 11 | 1 | PKC1_CARMO |

| | | | | | | |
|----|---|------|----|---|------------|---------------------|
| 34 | 3 | 10.7 | 11 | 1 | TKC2_CALVO | P41518 calliphora |
| 35 | 3 | 10.7 | 11 | 1 | TKN1_PSEGU | P42986 pseudophryn |
| 36 | 3 | 10.7 | 11 | 1 | TKN2_PSEGU | P42987 pseudophryn |
| 37 | 3 | 10.7 | 11 | 1 | TKN3_PSEGU | P42988 pseudophryn |
| 38 | 3 | 10.7 | 12 | 1 | SOI5_BACSU | P80863 bacillus su |
| 39 | 3 | 10.7 | 12 | 1 | TKN1_KASMA | P08611 kassina mac |
| 40 | 3 | 10.7 | 12 | 1 | TKN_KASSE | P08611 kassina sen |
| 41 | 3 | 10.7 | 13 | 1 | FIIB_CAVPO | P14445 cavia porce |
| 42 | 3 | 10.7 | 13 | 1 | FIIB_HYLLA | P14472 hylobates l |
| 43 | 3 | 10.7 | 13 | 1 | NEUT_BURMA | P81796 buru marlinu |
| 44 | 3 | 10.7 | 13 | 1 | NEUT_CHICK | P13724 gallus gall |
| 45 | 3 | 10.7 | 13 | 1 | NEUT_RANTE | P41536 rana tempor |

ALIGNMENTS

| RESULT | ID | NEUX_HUMAN | STANDARD | PRT | 9 AA |
|---------|---|--------------------------------|----------|--------|------|
| AC | P04277 | 20-MAR-1987 (Rel. 04, Created) | | | |
| DT | 20-MAR-1987 (Rel. 04, Last sequence update) | | | | |
| DT | 01-MAR-2002 (Rel. 41, Last annotation update) | | | | |
| DE | Neutensin-related peptide (NRP) (Kinensin). | | | | |
| OS | Homo sapiens (Human), and | | | | |
| OS | Bos taurus (Bovine), and | | | | |
| OS | Oryctolagus cuniculus (Rabbit). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| OX | NCBI_TaxID=9606, 9913, 9986; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE. | | | | |
| RC | SPECIES=Human; | | | | |
| RX | MEDLINE=86242180; PubMed=3087352; | | | | |
| RA | Mogard M.H., Kobayashi R., Chen C.F., Lee T.D., Reeve J.R. Jr., | | | | |
| RA | Shively J.E., Walsh J.H.; | | | | |
| RT | "The amino acid sequence of kinensin, a novel peptide isolated from | | | | |
| RT | pepsin-treated human plasma: homology with human serum albumin, | | | | |
| RT | neutensin and angiotensin."; | | | | |
| RT | Biochem. Biophys. Res. Commun. 136:983-988(1986). | | | | |
| RL | [2] | | | | |
| RN | SEQUENCE. | | | | |
| RP | SPECIES=Human, Bovine, and Rabbit; | | | | |
| RX | MEDLINE=87194805; PubMed=2437111; | | | | |
| RA | Carraway R.E., Mitra S.P., Cochrane D.E.; | | | | |
| RT | "Structure of a biologically active neutensin-related peptide | | | | |
| RT | obtained from pepsin-treated albumin(s)."; | | | | |
| RL | J. Biol. Chem. 262:5968-5973(1987). | | | | |
| CC | - FUNCTION: REGULATION OF FAT DIGESTION, LIPID ABSORPTION, AND | | | | |
| CC | BLOOD FLOW (POTENTIAL). | | | | |
| DR | PIR: A03239; ABHUSK. | | | | |
| DR | PIR: A26693; A26693. | | | | |
| KW | Hormone. | | | | |
| SO | SEQUENCE 9 AA; 1172 MW; C804DB4761F4140D CRC64; | | | | |
| QY | 19 ARRH 22 | | | | |
| DB | 2 ARRH 5 | | | | |
| QY | 14.3%; Score 4; DB 1; Length 9; | | | | |
| DB | Best Local Similarity 100.0%; Pred. No. 1e+05; | | | | |
| Matches | 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| RESULT | 2 | | | | |
| ID | NPA_BOVIN | STANDARD; | PRT; | 18 AA. | |
| AC | P15506; | | | | |
| DT | 01-APR-1990 (Rel. 14, Created) | | | | |
| DT | 01-APR-1990 (Rel. 14, Last sequence update) | | | | |
| DT | 01-APR-1990 (Rel. 14, Last annotation update) | | | | |

DE Neuropetide A.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_TaxID=9913;
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Brain;
 RX MEDLINE=86067985; PubMed=3865193;
 RA Yang H.-T., Fratta W., Majane E.A., Costa E.;
 RT Isolation, sequencing, synthesis, and pharmacological
 RT characterization of two brain neuropeptides that modulate the action
 RT of morphine.";
 RL Proc. Natl. Acad. Sci. U.S.A. 82:7757-7761(1985).
 CC -1- FUNCTION: MODULATES THE ACTION OF MORPHINE.
 DR PIR: A24749; A24749.
 KW Neuropeptide; Amidation.
 FT SIMILAR 15 18 TO NEUROPEPTIDE B (AA 5-8) (IDENTICAL).
 FT MOD_RES 18 18 AMIDATION.
 SQ SEQUENCE 18 AA; 1921 MW; EC52DAE1F45CF6B CRC64;

Query Match 14.3%; Score 4; DB 1; Length 18;
 Best Local Similarity 100.0%; Pred. No. 2.7e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GEG 7
 ||||
 DB 2 GEG 5

RESULT 3
 FIBB_FELCA STANDARD; PRT; 20 AA.
 AC P14469;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE.
 RA Blomback B., Blomback M., Grondahl N.J.;
 RT "Studies on fibrinopeptides from mammals.";
 RL Acta Chem. Scand. 19:1789-1791(1965).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_A6_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma.
 FT PEPTIDE 1 20 FIBRINOPEPTIDE B.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2328 MW; A829E393B8F627D0 CRC64;

Query Match 14.3%; Score 4; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 YDEG 26
 ||||

DB 5 YDEG 8

RESULT 4
 LYC_FELCA STANDARD; PRT; 20 AA.
 AC P37155;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1994 (Rel. 30, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Lysozyme C (EC 3.2.1.17) (1,4-beta-N-acetylmuramidase C) (Fragment).
 GN LYZ.
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.
 NCBI_TaxID=9685;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Milk;
 RX MEDLINE=90263403; PubMed=2344734;
 RA Halliday J.A., Bell K., McKenzie H.A., Shaw D.C.;
 RT "Feline whey proteins: identification, isolation and initial
 RT characterization of alpha-lactalbumin, beta-lactoglobulin and
 RT lysozyme.";
 RL Comp. Biochem. Physiol. 95B:773-779(1990).
 CC -1- FUNCTION: LYSOZYMES HAVE PRIMARILY BACTERIOLYTIC FUNCTION; THOSE
 CC IN TISSUES AND BODY FLUIDS ARE ASSOCIATED WITH THE MONOCYTE-
 CC MACROPHAGE SYSTEM AND ENHANCE THE ACTIVITY OF IMMUNOGENS.
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of the 1,4-beta-linkages between N-
 CC acetyl-D-glucosamine and N-acetylmuramic acid in peptidoglycan
 CC heteropolymers of the prokaryotes cell walls.
 CC -1- SUBUNIT: MONOMER.
 CC -1- MISCELLANEOUS: LYSOZYME C IS CAPABLE OF BOTH HYDROLYSIS AND
 CC TRANSGLYCOSYLATION. IT SHOWS ALSO A SLIGHT ESTERASE ACTIVITY. IT
 CC ACTS RAPIDLY ON BOTH PEPTIDE-SUBSTITUTED AND UNSUBSTITUTED
 CC PEPTIDOGLYCAN & SLOWLY, ON CHITIN OLIGOSACCHARIDES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 22 OF GLYCOSYL HYDROLASES.
 DR PIR: A60525; A60525.
 DR HSSP: P13376; 2EOL.
 DR InterPro: IPR001916; Lactalbum_lysozyme.
 DR Pfam: PF00062; Lys; 1.
 DR PROSITE: PS00128; LACTALBUMIN_LYSOZYME; PARTIAL.
 KW Hydrolyase; Glycosidase; Bacteriolytic enzyme; Milk.
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2314 MW; EB8824EA59425E13 CRC64;

Query Match 14.3%; Score 4; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RAEG 16
 ||||
 DB 13 RAEG 16

RESULT 5
 MIF_PIG STANDARD; PRT; 20 AA.
 AC P80928;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Macrophage migration inhibitory factor (MIF) (Phenylpyruvate
 DE tautomerase) (Glycosylation-inhibiting factor) (GIF) (Fragment).
 GN MIF.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE.
 RA Riviere S., Bouet F., Menez A., Galat A.;

RL Submitted (MAR-1997) to the SWISS-PROT data bank.
 CC -1- FUNCTION: THE EXPRESSION OF MIF AT SITES OF INFLAMMATION SUGGEST A
 CC ROLE FOR THE MEDIATOR IN REGULATING THE FUNCTION OF MACROPHAGE IN
 CC HOST DEFENSE. ALSO ACTS AS A PHENOLPYRUVATE TAUTOMERASE (BY
 CC SIMILARITY).
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE MIF FAMILY.
 DR HSP: P14174; IGIF.
 DR InterPro: IPR001398; MIF. PARTIAL.
 DR PROSITE: PS01158; MIF. PARTIAL.
 KM Isomerase; Macrophage; Inflammatory response; Cytokine.
 FT AC1_SITE 1 1 CATALYTIC BASE (BY SIMILARITY).
 FT NON_TER 20 20
 SQ SEQUENCE 20 AA; 2147 MW; 3517AF60F3012A61 CRC64;

Query Match 14.3%; Score 4; DB 1; Length 20;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 VPRA 14
 ||||
 DB 9 VPRA 12

RESULT 6
 FIBB_BISBO STANDARD; PRT; 21 AA.

ID FIBB_BISBO
 AC P14466;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Bison bonasus (European bison).
 OC Eumariyola; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bison.
 OX NCBI_TaxID=9902;
 RN [1]
 RP SEQUENCE.
 RA Blomback B., Blomback M., Grondahl N.J.;
 RT "Studies on fibrinopeptides from mammals.";
 RL Acta Chem. Scand. 19:1789-1791(1965).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2366 MW; 09EE75AF19E6363D CRC64;

Query Match 14.3%; Score 4; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 YDEG 26
 ||||
 DB 6 YDEG 9

RESULT 7
 FIBB_BUBBU

ID FIBB_BUBBU STANDARD; PRT; 21 AA.
 AC P14467;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Bubalus bubalis (domestic water buffalo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bubalus.
 OX NCBI_TaxID=89462;
 RN [1]
 RP SEQUENCE.
 RA Moss G.A., Doolittle R.F.;
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides.";
 RL Arch. Biochem. Biophys. 122:674-684(1967).
 RN [2]
 RP SEQUENCE.
 RC STRAIN-ITALIAN BREED;
 RX MEDLINE=76040091; PubMed=1180969;
 RA Balestrieri C., Colonna G., Irace G.;
 RT "Covalent structure of fibrinopeptides from buffaloes breeding in
 RT Italy.";
 RL Blochim. Biophys. Acta 405:517-521(1975).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 6 6 PYRROLIDONE CARBOXYLIC ACID.
 FT NON_TER 21 21
 SQ SEQUENCE 21 AA; 2379 MW; 09EE75BE4729163D CRC64;

Query Match 14.3%; Score 4; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 YDEG 26
 ||||
 DB 6 YDEG 9

RESULT 8
 FIBB_SYNCA STANDARD; PRT; 21 AA.

ID FIBB_SYNCA
 AC P14481;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Syncerus caffer (Cape buffalo).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Syncerus.
 OX NCBI_TaxID=9970;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=67209145; PubMed=6033721;
 RA Doolittle R.F., Schubert D., Schwartz S.A.;
 RT "Amino acid sequence studies on artiodactyl fibrinopeptides. I.
 RT Dromedary camel, mule deer, and cape buffalo.";

RL Arch. Biochem. Biophys. 118:456-467(1967).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 DR Interpro: IPR002181: fibrinogen.C.
 DR PROSITE: ps00514; FIBRINAG_C_DOMAIN; PARTIAL.
 KW Blood coagulation; Plasma; Sulfation.
 FT PEPTIDE 1 21 FIBRINOPEPTIDE B.
 FT MOD_RES 1 1 PYRROLIDONE CARBOXYLIC ACID.
 FT MOD_RES 6 6 SULFATION.
 FT NON_TER 21 21
 SO SEQUENCE 21 AA; 2353 MW; 09EF75AF033B863D CRC64;

Query Match 14.3%; Score 4; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 YDEG 26
 Db 6 YDEG 9

RESULT 9
 ID OMPL_ACTPL STANDARD; PRT; 21 AA.
 AC P80368;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 40 kDa major outer membrane protein (MOMP) (Fragment).
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Actinobacillus.
 OX NCBI_Taxid=715;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=598;
 RA MEDLINE=96065432; PubMed=7483902;
 RX Hartmann L., Schroeder W., Luebke-Becker A.;
 RT "Isolation of the major outer-membrane protein of Actinobacillus
 RT pleuropneumoniae and Haemophilus parasuis.";
 RL J. Vet. Med. B 42:59-63(1995).
 CC -1- FUNCTION: STRUCTURAL RIGIDITY OF THE OUTER MEMBRANE OF ELEMENTARY
 CC BODIES AND PORIN FORMING, PERMITTING DIFFUSION OF SOLUTES THROUGH
 CC THE INTRACELLULAR RETICULATE BODY MEMBRANE.
 CC -1- SUBUNIT: DISULFIDE BOND INTERACTIONS WITHIN AND BETWEEN MOMP
 CC MOLECULES & OTHER COMPONENTS FORM HIGH MOLECULAR-WEIGHT OLIGOMERS.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Outer membrane.
 KW Outer membrane; Transmembrane; Porin.
 FT NON_TER 21 21
 SO SEQUENCE 21 AA; 2293 MW; FFE7D12EA916563B CRC64;

Query Match 14.3%; Score 4; DB 1; Length 21;
 Best Local Similarity 100.0%; Pred. No. 3.1e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 14 AEGT 17
 Db 6 AEGT 9

RESULT 10
 IF2G_PIG
 ID IF2G_PIG STANDARD; PRT; 22 AA.
 AC P20461;

DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Eukaryotic translation initiation factor 2 subunit 3 (Eukaryotic
 DE translation initiation factor 2 gamma subunit) (eif-2-gamma)
 DE (Fragment).
 GN EIF2S3 OR EIF2G.
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RA SUZUKI H., MUKOYAMA E.B.;
 RT "pig liver translational initiation factor eif-2: N-terminal amino
 RT acid sequences of alpha and gamma subunits and the phosphorylation
 RT site structure.";
 RL Agric. Biol. Chem. 52:1397-1406(1988).
 CC -1- FUNCTION: EIF-2 FUNCTIONS IN THE EARLY STEPS OF PROTEIN SYNTHESIS
 CC BY FORMING A TERNARY COMPLEX WITH GTP AND INITIATOR TRNA. THIS
 CC COMPLEX BINDS TO A 40S RIBOSOMAL SUBUNIT, FOLLOWED BY MRNA BINDING
 CC TO FORM A 43S PREINITIATION COMPLEX. JUNCTION OF THE 60S RIBOSOMAL
 CC SUBUNIT TO FORM THE 80S INITIATION COMPLEX IS PRECEDED BY
 CC HYDROLYSIS OF THE GTP BOUND TO EIF-2 AND RELEASE OF AN EIF-2-GDP
 CC BINARY COMPLEX. IN ORDER FOR EIF-2 TO RECYCLE AND CATALYZE ANOTHER
 CC ROUND OF INITIATION, THE GDP BOUND TO EIF-2 MUST EXCHANGE WITH GTP
 CC BY WAY OF A REACTION CATALYZED BY EIF-2B.
 CC -1- SUBUNIT: HETEROTRIMER COMPOSED OF AN ALPHA, A BETA, AND A GAMMA
 CC CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE GTP-BINDING ELONGATION FACTOR FAMILY.
 CC EIF2G SUBFAMILY.
 DR PIR: PT0052; PT0052. Protein biosynthesis; GTP-binding.
 KW Initiation factor; Protein biosynthesis; GTP-binding.
 FT NON_TER 22 22
 SO SEQUENCE 22 AA; 2203 MW; 428BA7D7D18B03C CRC64;

Query Match 14.3%; Score 4; DB 1; Length 22;
 Best Local Similarity 100.0%; Pred. No. 3.2e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 Db 2 GTEA 5

RESULT 11
 ID CH60_MYCSM STANDARD; PRT; 28 AA.
 AC P80673;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE 60 kDa chaperonin (Protein Cpn60) (groEL protein) (Fragment).
 GN GROEL OR MOXA OR GROEL.
 GN Mycobacterium smegmatis.
 OS Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_Taxid=1772;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=ATCC 607 / MC(2)6 / NRRL B-692;
 RX MEDLIN=97387814; PubMed=9243799;
 RA Lundigan M.D., Arceneaux J.E.L., Zhu W., Byers B.R.;
 RT "Enhanced hydrogen peroxide sensitivity and altered stress protein
 RT expression in iron-starved Mycobacterium smegmatis.";
 RL Biometals 10:215-225(1997).
 CC -1- FUNCTION: PREVENTS MISFOLDING AND PROMOTES THE REFOLDING AND
 CC PROPER ASSEMBLY OF UNFOLDED POLYPEPTIDES GENERATED UNDER STRESS
 CC CONDITIONS.
 CC -1- SUBUNIT: OLIGOMER OF 14 SUBUNITS COMPOSED OF TWO STACKED RINGS OF
 CC 7 SUBUNITS (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE CHAPERONIN (HSP60) FAMILY.
 DR HSSP; P06139; IACN.
 DR InterPro: IPR001844; CHAPERONIN_CPN60.
 DR PROSITE: PS00296; CHAPERONIN_CPN60; PARTIAL.
 DR Chaperone; ATP-binding.
 FT NON_TER 28
 SQ SEQUENCE 28 AA; 3047 MW; 2F40F27B94EF8720 CRC64;

Query Match 14.3%; Score 4; DB 1; Length 28;
 Best Local Similarity 100.0%; Pred. No. 4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 18 EARR 21
 DB 9 EARR 12

RESULT 12
 TPIS_CANFA
 ID TPIS_CANFA STANDARD: PRT: 5 AA.
 AC P54714;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Triosephosphate isomerase (EC 5.3.1.1) (TIM) (Fragment).
 GN TP11.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE.

RC TISSUE-Heart;
 RX MEDLINE=98163340; PubMed=9504812;
 RA Dunn M.J., Corbett J.M., Wheeler C.H.;
 RT "HSC-2DPAGE and the two-dimensional gel electrophoresis database of dog heart proteins";
 RL Electrophoresis 18:2795-2802(1997).
 CC -1- CATALYTIC ACTIVITY: D-glyceraldehyde 3-phosphate = glyceralone phosphate.
 CC -1- PATHWAY: PLAYS AN IMPORTANT ROLE IN SEVERAL METABOLIC PATHWAYS.
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE TRIOSEPHOSPHATE ISOMERASE FAMILY.
 DR HSC-2DPAGE; P54714; DOG.
 DR InterPro: IPR000652; Trioseph. isomerase.
 DR PROSITE: PS00171; TIM; PARTIAL.
 KW Isomerase; Glycolysis; Gluconeogenesis; Fatty acid biosynthesis;
 KM Pentose shunt.
 FT NON_TER 1 1
 FT NON_TER 5 5
 SQ SEQUENCE 5 AA; 550 MW; 64444862C9A00000 CRC64;

Query Match 10.7%; Score 3; DB 1; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FVG 4
 DB 1 FVG 3

RESULT 13
 UPAL_HUMAN STANDARD: PRT: 8 AA.
 AC P30087;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Unknown protein from 2D-page of plasma (Spot 2) (Fragment).
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Plasma;
 RX MEDLINE=93092937; PubMed=1459097;
 RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
 RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
 RA Hochstrasser D.F.;
 RT "Plasma protein map: an update by microsequencing";
 RL Electrophoresis 13:707-714(1992).
 CC -1- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
 CC PROTEIN IS: 4.9, ITS MW IS: 65 kDa.
 DR SWISS-2DPAGE; P30087; HUMAN.
 FT NON_TER 1 1
 FT UNSURE 8 8
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 944 MW; C01772C455BB06DA CRC64;

Query Match 10.7%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 VPR 13
 DB 6 VPR 8

RESULT 14
 VGLG_HSV2B
 ID VGLG_HSV2B STANDARD: PRT: 8 AA.
 AC P81780;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Glycoprotein G (Fragment).
 OS Herpes simplex virus (type 2 / strain B43278R).
 CC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Alphaherpesvirinae; Simplexvirus.
 OX NCBI_TaxID=103921;
 RN [1]
 RP SEQUENCE.
 RA Liljeqvist J.-A., Svennerholm B., Bergstrom T.;
 RL Submitted (Apr-1999) to the SWISS-PROT data bank.
 CC -1- MISCELLANEOUS: THERE ARE SEVEN EXTERNAL GLYCOPROTEINS IN HSV1 AND
 CC 2: GH, GB, GC, GD, GI, AND GE.
 CC -1- MISCELLANEOUS: GLYCOPROTEIN G IS MUCH LARGER IN HSV-2 THAN IN
 CC HSV-1.
 KW Glycoprotein.
 FT NON_TER 8 8
 FT NON_TER 8 8
 SQ SEQUENCE 8 AA; 683 MW; 7B47686772C865B8 CRC64;

Query Match 10.7%; Score 3; DB 1; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 GVP 12
 DB 3 GVP 5

RESULT 15
 FIBB_PAPAN STANDARD: PRT: 9 AA.
 AC P19344;
 DT 01-NOV-1990 (Rel. 16, Created)
 DT 01-NOV-1990 (Rel. 16, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen beta chain [Contains: Fibrinopeptide B] (Fragment).
 GN FGB.
 OS Papio anubis (Olive baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Papio.
 OX NCBI_TaxID=9555;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=84161822; PubMed=6423621;
 RA Nakamura S., Takenaka O., Takahashi K.;
 RT "Fibrinopeptides A and B of baboons (Papio anubis, Papio hamadryas,
 and Theropithecus gelada): their amino acid sequences and
 evolutionary rates and a molecular phylogeny for the baboons.";
 RL J. Biochem. 94:1973-1978(1983).
 CC -1- FUNCTION: FIBRINOGEN HAS A DOUBLE FUNCTION: YIELDING MONOMERS THAT
 CC POLYMERIZE INTO FIBRIN AND ACTING AS A COFACTOR IN PLATELET
 CC AGGREGATION.
 CC -1- SUBUNIT: HEXAMER CONTAINING 2 SETS OF 3 NONIDENTICAL CHAINS
 CC (ALPHA, BETA AND GAMMA), LINKED TO EACH OTHER BY DISULFIDE BONDS.
 CC -1- MISCELLANEOUS: CONVERSION OF FIBRINOGEN TO FIBRIN IS TRIGGERED BY
 CC THROMBIN, WHICH CLEAVES FIBRINOPEPTIDES A AND B FROM ALPHA & BETA
 CC CHAINS, AND THUS EXPOSES THE N-TERMINAL POLYMERIZATION SITES
 CC RESPONSIBLE FOR THE FORMATION OF THE SOFT CLOT.
 CC PIR: D28854; D28854.
 DR InterPro: IPR002181; Fibrinogen_C.
 DR PROSITE: PS00514; FIBRIN_AG_C_DOMAIN; PARTIAL.
 KM Blood coagulation; Plasma.
 FT PEPTIDE 1 9 FIBRINOPEPTIDE B.
 FT NON_TER 9 9
 SQ SEQUENCE 9 AA; 1076 MW; DDF6409C7287B06 CRC64;

Query Match 10.7%; Score 3; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 1e+05;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EGI 7
 III
 Db 3 EGI 5

Search completed: June 26, 2002, 13:53:49
 Job time: 161 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:50:47 ; Search time 26.25 Seconds
(without alignments)
184.528 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28

Sequence: 1 DFVGEGLYGVPAEGTEARHRYDEGVR 28

Scoring table:

OLIGO
Gapop 60.0 , Gapept 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 12285

Minimum DB seq length: 0
Maximum DB seq length: 28

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 17.9 | 23 | 4 | Q9UC01 | Q9UC01 homo sapien |
| 2 | 17.9 | 26 | 12 | Q9ORV0 | Q9ORV0 hepatitis c |
| 3 | 17.9 | 28 | 12 | Q9QUT2 | Q9QUT2 tt virus. o |
| 4 | 14.3 | 13 | 8 | Q9T569 | Q9T569 zea mays (m |
| 5 | 14.3 | 14 | 2 | Q9R8J2 | Q9R8J2 chlamydia t |
| 6 | 14.3 | 14 | 2 | Q9R8J0 | Q9R8J0 chlamydia t |
| 7 | 14.3 | 14 | 2 | Q9R8I8 | Q9R8I8 chlamydia t |
| 8 | 14.3 | 14 | 2 | Q9R8I6 | Q9R8I6 chlamydia t |
| 9 | 14.3 | 14 | 2 | Q9R8I4 | Q9R8I4 chlamydia t |
| 10 | 14.3 | 14 | 2 | Q9R8I2 | Q9R8I2 chlamydia t |
| 11 | 14.3 | 14 | 2 | Q9R8I0 | Q9R8I0 chlamydia t |
| 12 | 14.3 | 14 | 2 | Q9R8H8 | Q9R8H8 chlamydia t |
| 13 | 14.3 | 14 | 2 | Q9R8H6 | Q9R8H6 chlamydia t |
| 14 | 14.3 | 14 | 2 | Q9R8H4 | Q9R8H4 chlamydia t |
| 15 | 14.3 | 14 | 2 | Q9R8H2 | Q9R8H2 chlamydia t |
| 16 | 14.3 | 14 | 2 | Q9R8H0 | Q9R8H0 chlamydia t |

| | | | | | | |
|----|---|------|----|----|--------|--------------------|
| 17 | 4 | 14.3 | 14 | 2 | Q9R8G8 | Q9R8G8 chlamydia t |
| 18 | 4 | 14.3 | 14 | 2 | Q9R8G6 | Q9R8G6 chlamydia t |
| 19 | 4 | 14.3 | 14 | 2 | Q9R8G4 | Q9R8G4 chlamydia t |
| 20 | 4 | 14.3 | 14 | 2 | Q9R8G3 | Q9R8G3 chlamydia t |
| 21 | 4 | 14.3 | 14 | 2 | Q9R8G1 | Q9R8G1 chlamydia t |
| 22 | 4 | 14.3 | 14 | 2 | Q9R8F9 | Q9R8F9 chlamydia t |
| 23 | 4 | 14.3 | 14 | 2 | Q9R8F7 | Q9R8F7 chlamydia t |
| 24 | 4 | 14.3 | 14 | 2 | Q9R8F5 | Q9R8F5 chlamydia t |
| 25 | 4 | 14.3 | 14 | 2 | Q9R8F2 | Q9R8F2 chlamydia t |
| 26 | 4 | 14.3 | 14 | 2 | Q9R8F0 | Q9R8F0 chlamydia t |
| 27 | 4 | 14.3 | 14 | 2 | Q9R8E9 | Q9R8E9 chlamydia t |
| 28 | 4 | 14.3 | 14 | 2 | Q9R8E8 | Q9R8E8 chlamydia t |
| 29 | 4 | 14.3 | 14 | 2 | Q9R8E7 | Q9R8E7 chlamydia t |
| 30 | 4 | 14.3 | 14 | 2 | Q9R8E6 | Q9R8E6 chlamydia t |
| 31 | 4 | 14.3 | 14 | 2 | Q9S3K8 | Q9S3K8 chlamydia t |
| 32 | 4 | 14.3 | 14 | 2 | Q9S3K6 | Q9S3K6 chlamydia t |
| 33 | 4 | 14.3 | 16 | 2 | Q9R4L1 | Q9R4L1 spiroplasma |
| 34 | 4 | 14.3 | 16 | 12 | Q04246 | Q04246 human para |
| 35 | 4 | 14.3 | 17 | 6 | Q95M99 | Q95M99 equus cabal |
| 36 | 4 | 14.3 | 17 | 13 | Q9PRU8 | Q9PRU8 gallus gall |
| 37 | 4 | 14.3 | 19 | 2 | Q31045 | Q31045 streptomyc |
| 38 | 4 | 14.3 | 19 | 8 | Q95G13 | Q95G13 abies alba |
| 39 | 4 | 14.3 | 19 | 11 | Q9QVNI | Q9QVNI rattus sp. |
| 40 | 4 | 14.3 | 20 | 4 | Q9N226 | Q9N226 homo sapien |
| 41 | 4 | 14.3 | 20 | 12 | Q83272 | Q83272 cucurbit mo |
| 42 | 4 | 14.3 | 20 | 13 | Q9PRR4 | Q9PRR4 scyllorhinu |
| 43 | 4 | 14.3 | 20 | 13 | Q9PRR3 | Q9PRR3 scyllorhinu |
| 44 | 4 | 14.3 | 22 | 10 | Q9S899 | Q9S899 pinus monti |
| 45 | 4 | 14.3 | 22 | 11 | Q9JHT2 | Q9JHT2 rattus norv |

ALIGNMENTS

| | | | | |
|---|-----------------|--------------------|-----------|------------|
| RESULT 1 | Q9UC01 | PRELIMINARY: | PRT: | 23 AA. |
| ID Q9UC01 | Q9UC01 | | | |
| AC Q9UC01 | Q9UC01 | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Created) | | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update) | | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update) | | | | |
| DE Y-23-R-NH2-PREDICTED INSULIN-LIKE GROWTH FACTOR IB- (103-124) EI AMIDE | | | | |
| DE IBEI. | | | | |
| OS Homo sapiens (Human) | | | | |
| OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| OX NCBI_TaxID=9606; | | | | |
| RN [1] | | | | |
| RP SEQUENCE FROM N.A. | | | | |
| RX MEDLINE=92390398; PubMed=1325646; | | | | |
| RA Siegfried J.M., Kasprzyk P.G., Treston A.M., Mulshine J.L., | | | | |
| RA Quinn R.A., Cuttitta F.; | | | | |
| RT "A mitogenic peptide amide encoded within the E peptide domain of the | | | | |
| RT insulin-like growth factor IB prohormone."; | | | | |
| RL Proc. Natl. Acad. Sci. U.S.A. 89:8107-8111(1992). | | | | |
| SO SEQUENCE 23 AA; 2508 MW; 40A5F3F520E3F789 CRC64; | | | | |
| Query Match | 17.9%; | Score 5; | DB 4; | Length 23; |
| Best Local Similarity | 100.0%; | Pred. No. 2.2e+02; | | |
| Matches 5; | Conservative 0; | Mismatches 0; | Indels 0; | Gaps 0; |
| QY 15 EGTEA 19 | | | | |
| DB 14 EGTEA 18 | | | | |
| RESULT 2 | Q9ORV0 | PRELIMINARY: | PRT: | 26 AA. |
| ID Q9ORV0 | Q9ORV0 | | | |
| AC Q9ORV0 | Q9ORV0 | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Created) | | | | |
| DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update) | | | | |

DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE E2 GLYCOPROTEIN HYPERVARIABLE REGION (FRAGMENT).
 OS Hepatitis C virus.
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
 OC Hepacivirus.
 OX NCBI_TaxID=11103;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-A-AS;
 RA Yeh C.-T.;
 RT "Replication of hepatitis C virus in the ascitic mononuclear cells and
 development of distinct quasispecies in the ascitic fluid."
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF109739; AAD51570.1; -
 FT NON_TER 1 1
 FT 26 26
 SQ SEQUENCE 26 AA; 2462 MW; 945C5E047695C5CC CRC64;

Query Match 17.9%; Score 5; DB 12; Length 26;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEAR 20
 DB 7 GTEAR 11

RESULT 3
 ID 090TJ2 PRELIMINARY; PRT; 28 AA.
 AC 090TJ2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAY-2000 (Tremblrel. 13, Last annotation update)
 DE ORF2 (FRAGMENT).
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-C278;
 RA Gerner P.;
 RT "TT virus infection in healthy children and in children with chronic
 hepatitis B or C."
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF188238; AAC04578.1; -
 FT NON_TER 28 28
 SQ SEQUENCE 28 AA; 2733 MW; ADEB94AF7C15B22C CRC64;

Query Match 17.9%; Score 5; DB 12; Length 28;
 Best Local Similarity 100.0%; Pred. No. 2.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 GVPRA 14
 DB 19 GVPRA 23

RESULT 4
 ID 09T569 PRELIMINARY; PRT; 13 AA.
 AC 09T569;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
 DE MITOCHONDRIAL PLASMID S-2 DNA, 5' END (FRAGMENT).
 OS Zea mays (Maize).
 OC Mitochondrion.
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
 OC Panicoideae; Andropogoneae; Zea.

OX NCBI_TaxID=4577;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=SEED;
 RA Traynor P.L., Levings C.S. III.;
 RT "Transcription of the S-2 maize mitochondrial plasmid."
 RL Plant Mol. Biol. 7:255-263(1986).
 DR EMBL; M16992; AAA70281.1; -
 KW Mitochondrion.
 FT NON_TER 13 13
 SQ SEQUENCE 13 AA; 1626 MW; 7B5E03673918A1F4 CRC64;

Query Match 14.3%; Score 4; DB 8; Length 13;
 Best Local Similarity 100.0%; Pred. No. 1.4e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 19 ARRH 22
 DB 4 ARRH 7

RESULT 5
 ID 09R8J2 PRELIMINARY; PRT; 14 AA.
 AC 09R8J2;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN-REF A;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
 trachomatis in trachoma endemic villages, Kongwa, Tanzania."
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
 DR EMBL; AF070309; AAC25280.1; -
 FT NON_TER 1 1
 FT 14 14
 SQ SEQUENCE 14 AA; 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 DB 10 GTEA 13

RESULT 6
 ID 09R8J0 PRELIMINARY; PRT; 14 AA.
 AC 09R8J0;
 DT 01-MAY-2000 (Tremblrel. 13, Created)
 DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 GN OMP-1.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=A;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia

RT trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF070313; AAC25284.1; -.

FT NON_TER 1 1

SO SEQUENCE 14 AA: 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
DB 10 GTEA 13

RESULT 7

ID 09R818 PRELIMINARY; PRT; 14 AA.

AC 09R818: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).

GN OMP-1.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RA Hsieh Y.-H., Bobo L.D.;

RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF070317; AAC25288.1; -.

FT NON_TER 1 1

SO SEQUENCE 14 AA: 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
DB 10 GTEA 13

RESULT 8

ID 09R816 PRELIMINARY; PRT; 14 AA.

AC 09R816: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).

GN OMP-1.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RA Hsieh Y.-H., Bobo L.D.;

RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF070321; AAC25292.1; -.

FT NON_TER 1 1

SO SEQUENCE 14 AA: 1405 MW; 14D7E247233CC4A0 CRC64;

SEQ SEQUENCE 14 AA: 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
DB 10 GTEA 13

RESULT 9

ID 09R814 PRELIMINARY; PRT; 14 AA.

AC 09R814: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).

GN OMP-1.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RA Hsieh Y.-H., Bobo L.D.;

RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF070325; AAC25296.1; -.

FT NON_TER 1 1

SO SEQUENCE 14 AA: 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
DB 10 GTEA 13

RESULT 10

ID 09R812 PRELIMINARY; PRT; 14 AA.

AC 09R812: 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).

GN OMP-1.

OS Chlamydia trachomatis.

OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.

OX NCBI_TaxID=813;

RN [1]

RP SEQUENCE FROM N.A.

RA Hsieh Y.-H., Bobo L.D.;

RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF070329; AAC25300.1; -.

FT NON_TER 1 1

SO SEQUENCE 14 AA: 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 ||||
 Db 10 GTEA 13

RESULT 11
 O9R810 PRELIMINARY; PRT; 14 AA.
 AC O9R810;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 OMP-1.
 GN Chlamydia trachomatis.
 OS Chlamydia trachomatis.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-REF A1B;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF070333; AAC25304.1; -.
 FT NON_TER 1 1
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 ||||
 Db 10 GTEA 13

RESULT 12
 O9R8H8 PRELIMINARY; PRT; 14 AA.
 AC O9R8H8;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 OMP-1.
 GN Chlamydia trachomatis.
 OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-REF A1C;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF070337; AAC25308.1; -.
 FT NON_TER 1 1
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 ||||
 Db 10 GTEA 13

RESULT 13
 O9R8H6 PRELIMINARY; PRT; 14 AA.
 AC O9R8H6;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 OMP-1.
 GN Chlamydia trachomatis.
 OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-REF A1D;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF070341; AAC25312.1; -.
 FT NON_TER 1 1
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 ||||
 Db 10 GTEA 13

RESULT 14
 O9R8H4 PRELIMINARY; PRT; 14 AA.
 AC O9R8H4;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
 OMP-1.
 GN Chlamydia trachomatis.
 OS Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=813;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-REF A1E;
 RA Hsieh Y.-H., Bobo L.D.;
 RT "Diversity of major outer membrane protein (omp-1) of Chlamydia trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF070345; AAC25316.1; -.
 FT NON_TER 1 1
 FT NON_TER 14 14
 SQ SEQUENCE 14 AA; 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
 Best Local Similarity 100.0%; Pred. No. 1.5e+03;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
 ||||
 Db 10 GTEA 13

RESULT 15
 O9R8H2 PRELIMINARY; PRT; 14 AA.
 ID O9R8H2

AC Q9R8H2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE MAJOR OUTER MEMBRANE PROTEIN (FRAGMENT).
GN omp-1.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-REF A1F;
RA Hsieh Y.-H., Bobo L.D.;
RT "Diversity of major outer membrane protein (omp-1) of Chlamydia
trachomatis in trachoma endemic villages, Kongwa, Tanzania.";
RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF070349; AAC25320.1; -.
FT NON_TER 1 1
FT NON_TER 14 14
SQ SEQUENCE 14 AA; 1405 MW; 14D7E247233CC4A0 CRC64;

Query Match 14.3%; Score 4; DB 2; Length 14;
Best Local Similarity 100.0%; Pred. No. 1.5e+03;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEA 19
|||
Db 10 GTEA 13

Search completed: June 26, 2002, 13:53:26
Job time: 159 sec

PAGE BLANK

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:43:07 ; Search time 31.66 Seconds

(without alignments)
98.233 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28
Sequence: 1 DFVGEGLYGGVPRAGEFTEARRHYDEGVR 28

Scoring table:
Gapop 60.0 , Gapext 60.0

Searched: 747574 seqs, 111073796 residues

Word size : 0

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A.Geneseq_032802:*

- 1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 15 | 53.6 | 371 | 22 | AAU69875 |
| 2 | 15 | 53.6 | 371 | 22 | AAU61230 |
| 3 | 15 | 53.6 | 371 | 22 | AAE01362 |
| 4 | 15 | 53.6 | 371 | 19 | AAW71869 |
| 5 | 15 | 53.6 | 553 | 19 | AAAB69385 |
| 6 | 15 | 53.6 | 553 | 21 | AAAB28527 |
| 7 | 15 | 53.6 | 553 | 21 | AAAB28527 |
| 8 | 15 | 53.6 | 553 | 22 | AAU69763 |
| 9 | 15 | 53.6 | 553 | 22 | AAU04961 |
| 10 | 15 | 53.6 | 553 | 22 | AAU01117 |
| 11 | 15 | 53.6 | 553 | 22 | AAU39002 |

| | | | | | | |
|----|----|------|------|----|-----------|---------------------|
| 12 | 15 | 53.6 | 553 | 22 | AAU62150 | Human p501S invent |
| 13 | 15 | 53.6 | 553 | 22 | AAU74800 | Prostate tumour an |
| 14 | 15 | 53.6 | 595 | 22 | AAU01318 | Alpha prepro-p501S |
| 15 | 15 | 53.6 | 1079 | 22 | AAU74830 | Prostate tumour an |
| 16 | 14 | 50.0 | 29 | 22 | AAU69832 | Human prostate pro |
| 17 | 14 | 50.0 | 29 | 22 | AAU01187 | Human prostate-spe |
| 18 | 14 | 50.0 | 29 | 22 | AAU69072 | Human prostate-spe |
| 19 | 13 | 46.4 | 25 | 22 | AAU69815 | Human prostate-spe |
| 20 | 13 | 46.4 | 25 | 22 | AAU01170 | Human prostate-spe |
| 21 | 13 | 46.4 | 25 | 22 | AAU69055 | Human prostate-spe |
| 22 | 13 | 46.4 | 27 | 20 | AAU85071 | Human prostate-deri |
| 23 | 13 | 46.4 | 27 | 21 | AAU82921 | Human prostate-deri |
| 24 | 13 | 46.4 | 27 | 22 | AAU69941 | Human prostate-rel |
| 25 | 13 | 46.4 | 27 | 22 | AAU04208 | Prostate-specific |
| 26 | 13 | 46.4 | 27 | 22 | AAU01296 | p501S, peptide 296- |
| 27 | 13 | 46.4 | 255 | 20 | AAU85068 | Protein encoded by |
| 28 | 13 | 46.4 | 255 | 21 | AAU82926 | Human prostate-rel |
| 29 | 13 | 46.4 | 255 | 22 | AAU04205 | Prostate-specific |
| 30 | 11 | 39.3 | 15 | 22 | AAU69813 | Human prostate pro |
| 31 | 11 | 39.3 | 15 | 22 | AAU01168 | Human prostate-spe |
| 32 | 11 | 39.3 | 15 | 22 | AAU69053 | Human prostate-spe |
| 33 | 7 | 25.0 | 179 | 22 | AAU63079 | Human prostate-spe |
| 34 | 7 | 25.0 | 1958 | 15 | AAU60620 | Drosophila melanog |
| 35 | 6 | 21.4 | 17 | 22 | AAU69814 | Protein from ORF2 |
| 36 | 6 | 21.4 | 17 | 22 | AAU01169 | Human prostate pro |
| 37 | 6 | 21.4 | 17 | 22 | AAU698054 | Human prostate-spe |
| 38 | 6 | 21.4 | 99 | 22 | AAU43901 | Propionibacterium |
| 39 | 6 | 21.4 | 99 | 22 | AAU65825 | Propionibacterium |
| 40 | 6 | 21.4 | 109 | 22 | AAU60594 | Novel human diagno |
| 41 | 6 | 21.4 | 121 | 22 | AAU66673 | Propionibacterium |
| 42 | 6 | 21.4 | 121 | 22 | AAU63928 | Human cancer assoc |
| 43 | 6 | 21.4 | 150 | 22 | AAU16915 | Human nervous syst |
| 44 | 6 | 21.4 | 150 | 22 | AAU4375 | Human polypeptide |
| 45 | 6 | 21.4 | 157 | 22 | AAU1929 | Human oxidoreducta |

ALIGNMENTS

| | |
|----------|---|
| RESULT 1 | |
| AAU69875 | AAU69875 standard; Protein: 371 AA. |
| ID | AAU69875; |
| AC | AAU69875; |
| XX | 30-JAN-2002 (first entry) |
| DT | |
| XX | Human prostate cDNA encoded protein #69. |
| DE | |
| XX | Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen. |
| XX | |
| OS | Homo sapiens. |
| PN | WO200173032-A2. |
| XX | |
| PD | 04-OCT-2001. |
| XX | |
| PF | 27-MAR-2001; 2001WO-US09919. |
| XX | |
| PR | 27-MAR-2000; 2000US-0536857. |
| PR | 09-MAY-2000; 2000US-0568100. |
| PR | 12-MAY-2000; 2000US-0570737. |
| PR | 13-JUN-2000; 2000US-0593793. |
| PR | 27-JUN-2000; 2000US-0605783. |
| PR | 10-AUG-2000; 2000US-0636215. |
| PR | 29-AUG-2000; 2000US-0651236. |
| PR | 06-SEP-2000; 2000US-0657279. |
| PR | 02-OCT-2000; 2000US-0679426. |
| PR | 10-OCT-2000; 2000US-0685166. |
| XX | |
| PA | (CORI-) CORIXA CORP. |
| XX | |
| PI | Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD; |

PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
XX
XX WPI: 2001-639232/73.
DR N-PSDB: AAS64038.
XX
PT New human prostate-specific polypeptides and polynucleotides useful for
PT the diagnosis and treatment of cancer, especially prostate cancer -
XX
XX Claim 2: Page 487-488; 579pp: English.
PS
XX The invention relates to isolated prostate-specific
CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
CC antibodies raised against the polypeptides (or antigenic epitopes
CC derived from them) and antigen-presenting cells expressing the
CC polypeptides. The antibodies are useful for detecting the presence of
CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
CC the antigen-presenting cells are useful for stimulating and/or expanding
CC T cells specific for a tumour protein, and for inhibiting the development
CC of cancer especially prostate cancer. Compositions comprising the
CC polynucleotide and/or polypeptide are useful for stimulating an immune
CC response, and for treating cancer. The oligonucleotide is useful for
CC detecting cancer. The present sequence is a prostate specific
CC polypeptide of the invention.
CC
SQ Sequence 371 AA;
XX

Query Match 53.6%; Score 15; DB 22; Length 371;
Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFVGEGLYGVRAE 15
|
Db 268 dfvgeglygvprae 282

RESULT 2
AAM01230
ID AAM01230 standard; Protein: 371 AA.
XX
AC AAM01330;
XX
DT 04-OCF-2001 (first entry)
XX
DE P553S splice variant P553S-14 amino acid #2.
XX
KW Human: prostate cancer; prostate-specific; diagnosis; vaccine;
KW cytosstatic; gene therapy; metastasis.
XX
OS Homo sapiens.
XX
PN WO200151633-A2.
XX
XX 19-JUL-2001.
PD
XX 16-JAN-2001; 2001WO-US01574.
PF
XX 14-JAN-2000; 2000US-0483672.
PR
XX (CORI-) CORIXA CORP.
PA
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Palos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAM;
PI Wang A, Meagher MJ;
XX
XX WPI: 2001-425873/45.
DR
XX
XX New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
PS Claim 2; Page 464-466; 543pp: English.

XX
CC The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAM01115 to
CC AAM01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
CC
SQ Sequence 371 AA;
XX

Query Match 53.6%; Score 15; DB 22; Length 371;
Best Local Similarity 100.0%; Pred. No. 6.3e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DFVGEGLYGVRAE 15
|
Db 268 dfvgeglygvprae 282

RESULT 3
AAE01362
ID AAE01362 standard; Protein: 371 AA.
XX
AC AAE01362;
XX
DT 18-JUL-2001 (first entry)
XX
DE Human gene 11 encoded secreted protein HMBAR14, SEQ ID NO:84.
XX
KW Human: secreted protein; proliferative disorder; cancer; tumour; asthma;
KW foetal abnormality; developmental abnormality; hemotopoletic disorder;
KW immune system disorder; AIDS; autoimmune disease; rheumatoid arthritis;
KW Parkinson's disease; cognitive disorder; schizophrenia; skin disorder;
KW psoriasis; sepsis; diabetes; atherosclerosis; cardiovascular disorder;
KW inflammation; neurological disorder; Alzheimer's disease; food additive;
KW angiogenic disorder; kidney disorder; gastrointestinal disorder; allergy;
KW pregnancy-related disorder; endocrine disorder; infection; wound healing;
KW cell culture; chemotaxis; vulnery; binding partner identification;
KW gene therapy.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX Peptide 1..48
XX FT /label= Signal_peptide
XX FT 49..371
XX FT /label= Mature_human_secreted_protein
XX FT MISC-difference 20
XX FT /label= Unknown
XX FT /note= "Encoded by TNC"
XX
XX WO200134629-A1.
XX
XX 17-MAY-2001.
PD
XX 08-NOV-2000; 2000WO-US30654.
PF
XX 12-NOV-1999; 99US-0164835.
PR 27-JUL-2000; 2000US-0221142.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX Ruben SM, Komatsoulis GA, Wei P, Fiscella M, Baker KP;
XX

| | | |
|----------|---|--|
| DR | WPI: 2001-308779/32. | |
| DR | N-PSDB: AAD05230. | |
| XX | | |
| PT | New nucleic acid encoding one of 21 human secreted proteins for | |
| PT | diagnosing, preventing, treating or ameliorating medical conditions, | |
| PT | such as autoimmune disease and cancer, and used as a food additive or | |
| XX | preservative - | |
| XX | | |
| PS | Claim 11; Page 438-440; 490pp; English. | |
| XX | | |
| CC | AAD05220-AAD05282 represent cDNAs corresponding to 21 human secreted | |
| CC | protein genes, and AAE01352-AAE01413 represent the proteins they encode. | |
| CC | AAE01415-AAE01433 represent human secreted protein fragments or variants. | |
| CC | The secreted proteins and their genes are useful for preventing, treating | |
| CC | or ameliorating medical conditions, e.g., by protein or gene therapy. | |
| CC | Pathological conditions can be diagnosed by determining the amount of the | |
| CC | new protein in a sample or by determining the presence of mutations in | |
| CC | the new genes. Specific uses are described for each of the 21 genes, | |
| CC | based on the tissues in which they are most highly expressed, and include | |
| CC | developing products for the diagnosis or treatment of proliferative | |
| CC | disorders, cancer, tumours, foetal and developmental abnormalities, | |
| CC | haematopoietic disorders, diseases of the immune system, AIDS, autoimmune | |
| CC | diseases (e.g., rheumatoid arthritis), inflammation, allergies, | |
| CC | neurological disorders (e.g., Alzheimer's disease, Parkinson's disease), | |
| CC | conjunctive disorders, schizophrenia, asthma, skin disorders (e.g., | |
| CC | psoriasis), sepsis, diabetes, atherosclerosis, cardiovascular disorders, | |
| CC | angiogenic disorders, kidney disorders, gastrointestinal disorders, | |
| CC | pregnancy-related disorders, endocrine disorders, and infections. The | |
| CC | proteins can also be used to aid wound healing and epithelial cell | |
| CC | proliferation, to prevent skin aging due to sunburn, to maintain organs | |
| CC | before transplantation, for supporting cell culture of primary tissues, | |
| CC | to regenerate tissues, to identify their cognate ligands or binding | |
| CC | partners, and in chemotaxis, and can be used as a food additive or a | |
| CC | preservative to modify storage properties. Antibodies specific for a | |
| CC | protein of the invention can be used in alleviating symptoms associated | |
| CC | with the disorders mentioned above, and in diagnostic immunoassays e.g., | |
| CC | radioimmunoassay or enzyme linked immunosorbent assay (ELISA). | |
| CC | The present sequence represents a human secreted protein of | |
| CC | the invention. | |
| XX | | |
| SQ | Sequence 371 AA; | |
| QY | 1 DEVGGLTQGVPRAE 15 | 53.6%; Score 15; DB 22; Length 371; |
| | | Best Local Similarity 100.0%; Pred. No. 6.3e-07; |
| | Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | |
| DB | 268 dtvgegllyqgvprae 282 | |
| RESULT 4 | | |
| ID | AAW71869 | |
| XX | AAW71869 standard; Protein: 553 AA. | |
| XX | | |
| AC | AAW71869; | |
| XX | | |
| DT | 06-JAN-1999 (first entry) | |
| XX | | |
| DE | Amino acid encoded by prostate tumour clone LI-12. | |
| XX | | |
| XX | Prostate; cancer; tumour; vaccine; immunogen; clone. | |
| OS | Homo sapiens. | |
| XX | | |
| PN | W09837093-A2. | |
| XX | | |
| PD | 27-AUG-1998. | |
| XX | | |
| PF | 25-FEB-1998; 98WO-US03492. | |
| XX | | |
| PR | 09-FEB-1998; 98US-0020956. | |

```

PR    25-FEB-1997;   97US-0806099.
PR    01-AUG-1997;   97US-0904804.
XX
PA    (CORI-) CORIXA CORP.
XX
PI    Dillon DC, Xu J;
XX
DR    WPI: 1998-609886/51.
XX    N-PSDB: AAV61201.
XX
PT    Polypeptides comprising immunogenic portions of prostate proteins -
PS    used in a vaccine for the treatment of prostate cancer
XX
XX    Example 1: Page 82-84; 130pp; English.
XX
CC    The present sequence is an immunogenic portion of a prostate tumour
CC    protein. The immunogen, or the DNA encoding it, can be used as a
CC    vaccine for the treatment of prostate cancer. The immunogen was
CC    isolated from a prostate tumour cDNA library obtained by subtracting
CC    a prostate tumour cDNA expression library with a normal tissue cDNA
CC    library.
XX
SQ    Sequence      553 AA;

Query Match          53.6%; Score 15; DB 19; Length 553;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 DFVGEGLYGVPRAE 15
        |||||||
Db       294 dfvgegllygvprrae 308

RESULT      5
AAW69385
ID      AAW69385 standard; Protein: 553 AA.
XX
AC      AAW69385;
XX
DT      08-DEC-1998 (first entry)
XX
DE      Prostate tumour specific gene clone L1-12 protein.
XX
KW      Prostate tumour specific gene; human; prostate cancer; detection;
KW      therapy.
XX
OS      Homo sapiens.
XX
PN      WO9837418-A2.
XX
PD      27-AUG-1998.
XX
PE      25-FEB-1998;   98WO-US03690.
XX
PR      09-FEB-1998;   98US-0904809.
PR      25-FEB-1997;   97US-0806596.
PR      01-AUG-1997;   97US-0904809.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Dillon DC, Xu J;
XX
DR      WPI: 1998-480805/41.
DR      N-PSDB: AAV58586.
XX
PT      Novel human prostate specific tumour protein and fragments - useful
PT      for detecting and treating prostate cancers
XX
XX    Example 1: Page 87-89; 141pp; English.
XX
CC    This sequence is encoded by a human prostate tumour specific gene, and
CC    can be used in the method of the invention. The method is for detecting

```

CC prostate cancer comprises contacting a biological sample with an agent
CC able to bind an immunogenic portion of a prostate protein (such as
CC this protein sequence). An antibody which binds to an immunogenic
CC portion of the prostate protein, and the method can be used to detect,
CC monitor progression of, or treat prostate cancers. The antibody may
CC also be conjugated to a therapeutic agent for use in therapy of prostate
CC cancers.

XX
SQ Sequence 553 AA;

Query Match 53.6%; Score 15; DB 19; Length 553;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
|
Db 294 dfvgegilygvprae 308

RESULT 6

AAB28527
ID AAB28527 standard; Protein: 553 AA.

AC AAB28527;

DT 07-FEB-2001 (first entry)

DE Protein encoded by human breast tumour CDNA clone p5015.

KW Human; breast tumour antigen; cytostatic; immunotherapy;

KM breast cancer; vaccine.

OS Homo sapiens.

PN WO200061756-A2.

PD 19-OCT-2000.

PF 10-APR-2000; 2000WO-US09688.

PR 09-APR-1999; 99US-0288950.

PR 02-JUL-1999; 99US-0346327.

PA (CORI-) CORIXA CORP.

PI Reed SG, Xu J, Dillon DC;

DR WPI; 2000-638568/61.

DR N-PSDB; AAC79473.

PT A novel isolated polypeptide comprising an immunogenic portion of a
PT breast cancer protein useful in the detection and treatment of breast
PT cancer -

PS Claim 2; Page 92-93; 95pp; English.

XX The present sequence is encoded by a cDNA sequence which was isolated
CC from a breast tumour cDNA library. It is provided in a specification
CC relating to compounds for immunotherapy and diagnosis of breast cancer.
CC Breast tumour antigens and the polynucleotides that encode them may be
CC used in the production of a pharmaceutical composition to be used in the
CC treatment of breast cancer. Proliferated T cells and incubated antigen
CC presenting cells are also required. The polypeptides and polynucleotides
CC may also be used to produce a vaccine.

XX Sequence 553 AA;

Query Match 53.6%; Score 15; DB 21; Length 553;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
|
Db 294 dfvgegilygvprae 308

RESULT 7

AAV82002
ID AAV82002 standard; Protein: 553 AA.

AC AAV82002;

DT 13-JUN-2000 (first entry)

DE Human immunogenic prostate tumour protein sequence SEQ ID NO:113.

KW Human; prostate cancer; diagnosis; tumour; gene therapy; detection;

KM immunogenic; cytostatic; vaccine.

OS Homo sapiens.

PN WO200004149-A2.

PD 27-JAN-2000.

PF 14-JUL-1999; 99WO-US15838.

PR 14-JUL-1998; 98US-0115453.

PR 14-JUL-1998; 98US-0116134.

PR 23-SEP-1998; 98US-0159812.

PR 15-SEP-1998; 98US-0159812.

PR 15-JAN-1999; 99US-0232149.

PR 15-JAN-1999; 99US-0232880.

PR 09-APR-1999; 99US-0288946.

PA (CORI-) CORIXA CORP.

PI Dillon DC, Harlocker SL, Yugu J, Xu J, Mitcham JL;

DR WPI; 2000-171268/15.

PT New polypeptide useful for treating and diagnosing prostate cancer
PT comprises an immunogenic portion of prostate tumor protein -

PS Claim 3; Page 138-139; 263pp; English.

XX The present invention describes isolated polypeptides, comprising an
CC immunogenic portion of a prostate tumour protein (PTP). The polypeptides
CC and polynucleotides encoding them have cytostatic activity and can be
CC used in vaccines and in gene therapy. The polypeptides and
CC polynucleotides encoding them, antigen presenting cells which express
CC the polypeptides, antibodies against the polypeptides and vaccines
CC comprising them can be used for inhibiting the development of prostate
CC cancer in a patient. The polypeptides can be used to generate antibodies
CC or anti-idiotypic antibodies for passive immuno therapy. A portion of
CC the polynucleotides encoding the polypeptides can be used as a probe or
CC to modulate the expression of the polypeptides. AAA06241 to AAA06691 and
CC AAV82000 to AAV82020 represent sequences used in the exemplification of
CC the present invention.

XX Sequence 553 AA;

Query Match 53.6%; Score 15; DB 21; Length 553;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
|
Db 294 dfvgegilygvprae 308

RESULT 8

AAU69763

GenCore version 4.5.
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 12:36:27 ; Search time 67.61 Seconds
(without alignments)
199.783 Million cell updates/sec

Title: US-09-838-785-2
Perfect score: 2861
Sequence: 1 MVQRLWVSRLLRHRKAQLL.....AIYFATQVVFDSKLAKYSA 553

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2.6/prodata/2/iaa/5A_COMB.pep.*
2: /cgn2.6/prodata/2/iaa/5B_COMB.pep.*
3: /cgn2.6/prodata/2/iaa/6A_COMB.pep.*
4: /cgn2.6/prodata/2/iaa/6B_COMB.pep.*
5: /cgn2.6/prodata/2/iaa/6C_COMB.pep.*
6: /cgn2.6/prodata/2/iaa/6D_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 2861 | 100.0 | 553 | 4 | US-09-020-956-113 |
| 2 | 2861 | 100.0 | 553 | 4 | US-09-030-607-113 |
| 3 | 2861 | 100.0 | 553 | 4 | US-09-439-313-113 |
| 4 | 1287 | 45.0 | 255 | 4 | US-09-071-710-36 |
| 5 | 1287 | 45.0 | 255 | 4 | US-09-525-397-36 |
| 6 | 452 | 15.8 | 84 | 4 | US-09-439-313-571 |
| 7 | 330 | 11.5 | 516 | 1 | US-08-356-340-4 |
| 8 | 330 | 11.5 | 516 | 2 | US-08-786-555-4 |
| 9 | 324.5 | 11.3 | 525 | 1 | US-08-356-340-2 |
| 10 | 324.5 | 11.3 | 525 | 2 | US-08-786-555-2 |
| 11 | 312 | 10.9 | 58 | 4 | US-09-439-313-547 |
| 12 | 304 | 10.6 | 56 | 4 | US-09-439-313-564 |
| 13 | 243 | 8.5 | 44 | 4 | US-09-071-710-37 |
| 14 | 243 | 8.5 | 44 | 4 | US-09-525-397-37 |
| 15 | 157 | 5.5 | 29 | 4 | US-09-439-313-546 |
| 16 | 146 | 5.1 | 27 | 4 | US-09-071-710-39 |
| 17 | 146 | 5.1 | 27 | 4 | US-09-525-397-39 |
| 18 | 146 | 5.1 | 27 | 4 | US-09-439-313-566 |
| 19 | 137 | 4.8 | 25 | 4 | US-09-439-313-520 |
| 20 | 135 | 4.7 | 27 | 4 | US-09-071-710-38 |
| 21 | 135 | 4.7 | 27 | 4 | US-09-525-397-38 |
| 22 | 120 | 4.2 | 21 | 4 | US-09-439-313-496 |
| 23 | 120 | 4.2 | 21 | 4 | US-09-439-313-521 |
| 24 | 119 | 4.2 | 22 | 4 | US-09-439-313-565 |
| 25 | 115 | 4.0 | 22 | 4 | US-09-439-313-559 |
| 26 | 114 | 4.0 | 22 | 4 | US-09-439-313-557 |
| 27 | 114 | 4.0 | 22 | 4 | US-09-439-313-561 |

| | | | | | | |
|----|-------|-----|-----|---|-------------------|--------------------|
| 28 | 112 | 3.9 | 20 | 4 | US-09-439-313-495 | Sequence 495, App |
| 29 | 110 | 3.8 | 427 | 4 | US-08-900-230-4 | Sequence 4, Appli |
| 30 | 109.5 | 3.8 | 455 | 1 | US-08-035-928-2 | Sequence 2, Appli |
| 31 | 109 | 3.8 | 22 | 4 | US-09-439-313-555 | Sequence 555, App |
| 32 | 107 | 3.7 | 20 | 4 | US-09-439-313-493 | Sequence 493, App |
| 33 | 107 | 3.7 | 20 | 4 | US-09-439-313-497 | Sequence 497, App |
| 34 | 106.5 | 3.7 | 436 | 6 | 5432081-10 | Patent No. 5432081 |
| 35 | 106 | 3.7 | 22 | 4 | US-09-439-313-569 | Sequence 569, App |
| 36 | 106 | 3.7 | 427 | 4 | US-09-199-737-4 | Sequence 4, Appli |
| 37 | 105 | 3.7 | 22 | 4 | US-09-439-313-553 | Sequence 553, App |
| 38 | 105 | 3.7 | 22 | 4 | US-09-439-313-572 | Sequence 572, App |
| 39 | 101 | 3.5 | 18 | 4 | US-09-439-313-548 | Sequence 548, App |
| 40 | 100 | 3.5 | 20 | 4 | US-09-439-313-490 | Sequence 490, App |
| 41 | 100 | 3.5 | 20 | 4 | US-09-439-313-499 | Sequence 499, App |
| 42 | 99 | 3.5 | 20 | 4 | US-09-439-313-563 | Sequence 563, App |
| 43 | 99 | 3.5 | 20 | 4 | US-09-439-313-570 | Sequence 570, App |
| 44 | 98 | 3.4 | 20 | 4 | US-09-439-313-498 | Sequence 498, App |
| 45 | 98 | 3.4 | 713 | 3 | US-09-335-409-11 | Sequence 11, Appli |

ALIGNMENTS

RESULT 1
US-09-020-956-113
; Sequence 113, Application US/09020956
; Patent No. 6261562
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SERD and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; US-09-020-956-113

Query Match 100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 3.7e-271; Indels 0; Gaps 0;
Matches 553; Conservative 0; Mismatches 0;
QY 1 MVQRLWVSRLLRHRKAQLLVNLLAFGLVCLAAAGTTVPPLLLVGVGVEKEKFTWVLGIG 60
|||||

THIS PAGE BLANK (USP'10)

```

Db      1  MQRIMVSLRLHRRKAOQLLVNLFEGLEVCLAAGITVPPILLLEVGEVEKFMVAVLGIG 60
Oy      61  PVLGLVCPVPLIGSASDHMRGRYGRRRPFTWALSIGLILSLFLIPRAGMLAGLCPDPRPL 120
Db      61  PVLGLVCPVPLIGSASDHMRGRYGRRRPFTWALSIGLILSLFLIPRAGMLAGLCPDPRPL 120
Oy      121  ETALLILGVGLDFCGOVCFTPLEALLSDLFDPDHCROAVSVAFMISLGGCLGYLLPA 180
Db      121  ETALLILGVGLDFCGOVCFTPLEALLSDLFDPDHCROAVSVAFMISLGGCLGYLLPA 180
Oy      181  IMDTSAALAPYLGTOBECLFGLTLFLTCVNAATLLVAEALGPTEPAEGISAPLSPH 240
Db      181  IMDTSAALAPYLGTOBECLFGLTLFLTCVNAATLLVAEALGPTEPAEGISAPLSPH 240
Oy      241  CCPCRARLAFNRNIGALLPRLHOLCCMRPRTLRLFVAELCSNMALMTFTLYTDVYGBGL 300
Db      241  CCPCRARLAFNRNIGALLPRLHOLCCMRPRTLRLFVAELCSNMALMTFTLYTDVYGBGL 300
Oy      301  YOGVPAEPGTETARRHYDEGVNMGSLGLFLQCAISLVFSVMORLYOFGTRAVYLA5VA 360
Db      301  YOGVPAEPGTETARRHYDEGVNMGSLGLFLQCAISLVFSVMORLYOFGTRAVYLA5VA 360
Oy      361  AFPVAAAGATCLSHSVAVVYASALTGTFTSALQILPYTLASLYHREKOVFLPKYRGDTG 420
Db      361  AFPVAAAGATCLSHSVAVVYASALTGTFTSALQILPYTLASLYHREKOVFLPKYRGDTG 420
Oy      421  ASSEDSIMTSFLPGPKGAPFPNGHVAGAGSGLLPPPALCASACDVSVRVVGEPTEA 480
Db      421  ASSEDSIMTSFLPGPKGAPFPNGHVAGAGSGLLPPPALCASACDVSVRVVGEPTEA 480
Oy      481  RVPVGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSOSVTAYMVSAGLGLVATYFATQ 540
Db      481  RVPVGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSOSVTAYMVSAGLGLVATYFATQ 540
Oy      541  VVFDKSDLAKYSA 553
Db      541  VVFDKSDLAKYSA 553

RESULT 2
US-09-030-607-113
; Sequence 113, Application US/09030607
; Patent No. 6262245
; GENERAL INFORMATION:
; APPLICANT: Xu, JIANGCHUN
; APPLICANT: DILLON, David C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS FO
; NUMBER OF SEQUENCES: 224
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED AND BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/030,607
; FILING DATE: 25-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427G3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:

```

```

; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGIN: SOURCE:
; ORGANISM: Homo sapiens
; US-09-030-607-113

Query Match      100.0%; Score 2861; DB 4; Length 553;
Best Local Similarity 100.0%; Pred. No. 3,7e-271;
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      1  MQRIMVSLRLHRRKAOQLLVNLFEGLEVCLAAGITVPPILLLEVGEVEKFMVAVLGIG 60
Db      1  MQRIMVSLRLHRRKAOQLLVNLFEGLEVCLAAGITVPPILLLEVGEVEKFMVAVLGIG 60
Oy      61  PVLGLVCPVPLIGSASDHMRGRYGRRRPFTWALSIGLILSLFLIPRAGMLAGLCPDPRPL 120
Db      61  PVLGLVCPVPLIGSASDHMRGRYGRRRPFTWALSIGLILSLFLIPRAGMLAGLCPDPRPL 120
Oy      121  ETALLILGVGLDFCGOVCFTPLEALLSDLFDPDHCROAVSVAFMISLGGCLGYLLPA 180
Db      121  ETALLILGVGLDFCGOVCFTPLEALLSDLFDPDHCROAVSVAFMISLGGCLGYLLPA 180
Oy      181  IMDTSAALAPYLGTOBECLFGLTLFLTCVNAATLLVAEALGPTEPAEGISAPLSPH 240
Db      181  IMDTSAALAPYLGTOBECLFGLTLFLTCVNAATLLVAEALGPTEPAEGISAPLSPH 240
Oy      241  CCPCRARLAFNRNIGALLPRLHOLCCMRPRTLRLFVAELCSNMALMTFTLYTDVYGBGL 300
Db      241  CCPCRARLAFNRNIGALLPRLHOLCCMRPRTLRLFVAELCSNMALMTFTLYTDVYGBGL 300
Oy      301  YOGVPAEPGTETARRHYDEGVNMGSLGLFLQCAISLVFSVMORLYOFGTRAVYLA5VA 360
Db      301  YOGVPAEPGTETARRHYDEGVNMGSLGLFLQCAISLVFSVMORLYOFGTRAVYLA5VA 360
Oy      361  AFPVAAAGATCLSHSVAVVYASALTGTFTSALQILPYTLASLYHREKOVFLPKYRGDTG 420
Db      361  AFPVAAAGATCLSHSVAVVYASALTGTFTSALQILPYTLASLYHREKOVFLPKYRGDTG 420
Oy      421  ASSEDSIMTSFLPGPKGAPFPNGHVAGAGSGLLPPPALCASACDVSVRVVGEPTEA 480
Db      421  ASSEDSIMTSFLPGPKGAPFPNGHVAGAGSGLLPPPALCASACDVSVRVVGEPTEA 480
Oy      481  RVPVGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSOSVTAYMVSAGLGLVATYFATQ 540
Db      481  RVPVGRGICLDLAILDSAFLLSQVAPSLFMGSIVQLSOSVTAYMVSAGLGLVATYFATQ 540
Oy      541  VVFDKSDLAKYSA 553
Db      541  VVFDKSDLAKYSA 553

RESULT 3
US-09-439-313-113
; Sequence 113, Application US/09439313
; Patent No. 6329505
; GENERAL INFORMATION:
; APPLICANT: Xu, JIANGCHUN
; APPLICANT: DILLON, David C.
; APPLICANT: MITCHELL, Jennifer L.
; APPLICANT: HARLOCKER, Susan Louise
; APPLICANT: JIANG YUQU
; APPLICANT: REED, Steven G.
; APPLICANT: KALOS, Michael
; APPLICANT: FANGER, Gary
; APPLICANT: RETTER, Mark
; APPLICANT: SOLK, John
; APPLICANT: DAY, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; DIAGNOSIS OF PROSTATE CANCER

```

THIS PAGE BLANK (USPTO)

ID AA069763 standard; Protein: 553 AA.
 XX
 AC AA069763;
 XX
 DT 30-JAN-2002 (first entry)
 XX
 DE Human prostate CDNA encoded protein #3.
 XX
 DE Human prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX
 KW Human; prostate cancer; cytostatic; immunostimulant; tumour; immunogen.
 XX
 OS Homo sapiens.
 XX
 PN WO200173032-A2.
 XX
 PD 04-OCT-2001.
 XX
 PF 27-MAR-2001; 2001WO-US09919.
 XX
 PR 27-MAR-2000; 2000US-0536857.
 PR 09-MAY-2000; 2000US-0568100.
 PR 12-MAY-2000; 2000US-0570737.
 PR 13-JUN-2000; 2000US-0593793.
 PR 27-JUN-2000; 2000US-0605783.
 PR 10-AUG-2000; 2000US-0636215.
 PR 29-AUG-2000; 2000US-0651236.
 PR 06-SEP-2000; 2000US-0651279.
 PR 02-OCT-2000; 2000US-0679426.
 PR 10-OCT-2000; 2000US-0685166.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Kalos MD;
 PI Fanger GR, Retter MW, Stolk JA, Day CH, Vedvick TS, Carter D;
 PI Li SX, Wang A, Skeiky YAM, Hepler WT, Henderson RA;
 XX
 DR WPI: 2001-639232/73.
 DR N-PSDB: AAS63557.
 XX
 PT New human prostate-specific polypeptides and polynucleotides useful for
 PT the diagnosis and treatment of cancer, especially prostate cancer -
 XX
 PS Claim 2; Page 269-270; 579pp; English.
 XX
 CC The invention relates to isolated prostate-specific
 CC polynucleotides, polypeptides, fusion proteins of the polypeptides,
 CC antibodies raised against the polypeptides (or antigenic epitopes
 CC derived from them) and antigen-presenting cells expressing the
 CC polypeptides. The antibodies are useful for detecting the presence of
 CC cancer, especially prostate cancer. The polypeptides, polynucleotides and
 CC the antigen-presenting cells are useful for stimulating and/or expanding
 CC T cells specific for a tumour protein, and for inhibiting the development
 CC of cancer especially prostate cancer. Compositions comprising the
 CC polynucleotide and/or polypeptide are useful for stimulating an immune
 CC response, and for treating cancer. The oligonucleotide is useful for
 CC detecting cancer. The present sequence is a prostate specific
 CC polypeptide of the invention.
 XX
 SQ Sequence 553 AA:

Query Match 53.6%; Score 15; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 8.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
 ||||||||||||
 DB 294 dfvgeglygvprae 308

RESULT 9
 AA004961
 ID AA004961 standard; Protein: 553 AA.
 XX

AC AA004961;
 XX
 DT 24-OCT-2001 (first entry)
 XX
 DE Human prostate tumour protein LI-12.
 XX
 DE Human prostate tumour protein; prostate cancer.
 XX
 KW Human; prostate tumour protein; prostate cancer.
 XX
 OS Homo sapiens.
 XX
 PN US6262245-B1.
 XX
 PD 17-JUL-2001.
 XX
 PF 25-FEB-1998; 98US-0030607.
 XX
 PR 25-FEB-1997; 97US-0806099.
 PR 01-AUG-1997; 97US-0904804.
 PR 09-FEB-1998; 98US-0020956.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Dillon DC;
 XX
 DR WPI: 2001-440862/47.
 DR N-PSDB: AAS10108.
 XX
 PT Novel polynucleotide encoding polypeptide comprising a portion of
 PT prostate tumour protein useful for inhibiting development of prostate
 PT cancer or for treating prostate cancer in a patient -
 XX
 PS Example 1; Column 125-127; 105pp; English.
 XX
 CC The sequence is a partial prostate tumour protein, encoded by a prostate
 CC tumour specific CDNA. The DNA is useful for inhibiting the development
 CC of prostate cancer or for treating prostate cancer in a patient.
 XX
 SQ Sequence 553 AA:

Query Match 53.6%; Score 15; DB 22; Length 553;
 Best Local Similarity 100.0%; Pred. No. 8.8e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
 ||||||||||||
 DB 294 dfvgeglygvprae 308

RESULT 10
 AA001117
 ID AA001117 standard; Protein: 553 AA.
 XX
 AC AA001117;
 XX
 DT 04-OCT-2001 (first entry)
 XX
 DE Human prostate-specific amino acid sequence LI-12.
 XX
 DE Human prostate cancer; prostate-specific; diagnosis; vaccine;
 KW Human; prostate cancer; prostate-specific; diagnosis; vaccine;
 KW cytostatic; gene therapy; metastasis.
 XX
 OS Homo sapiens.
 XX
 PN WO200151633-A2.
 XX
 PD 19-JUL-2001.
 XX
 PF 16-JAN-2001; 2001WO-US01574.
 XX
 PR 14-JAN-2000; 2000US-0483672.
 XX
 PA (CORI-) CORIXA CORP.

XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Fanger GR, Day CH, Retter MW, Stolk JA, Skeiky YAW;
PI Wang A, Meagher MJ;
DR WPI: 2001-425873/45.
XX
XX New polynucleotide encoding a prostate-specific protein, for
PT diagnosing, monitoring and treating prostate cancer in a patient and
PT for use in vaccines -
XX
XX Claim 2; Page 267-268; 543pp; English.
XX
XX The present invention describes polynucleotide sequences (I) which encode
CC prostate-specific proteins (II). (I) and (II) have cytostatic activity,
CC and can be used in vaccine production and gene therapy. (I), (II),
CC antibodies to (II), fusion proteins comprising (II), and isolated
CC T cells prepared using (I) or (II) are used to treat cancer in a patient.
CC (I) and the antibodies are also used in the detection of cancer in a
CC patient. The cancer that is diagnosed or treated is particularly
CC prostate cancer. (I) and (II) can be used in vaccines. The antibodies or
CC (I) can be used for monitoring the progression of cancer in a patient.
CC (I) and (II) can also be used to improve diagnostic and therapeutic
CC methods for prostate cancer. They can indicate the level of metastasis
CC as well as the prostate volume. AAH93357 to AAH93944 and AAH01115 to
CC AAH01318 represent polynucleotide and amino acid sequences used in the
CC exemplification of the present invention.
XX
XX Sequence 553 AA;

Query Match 53.6%; Score 15; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
|||||
Db 294 dfvgeglygvprae 308

RESULT 11
AAG99002
ID AAG99002 standard; Protein; 553 AA.
XX
XX AAG99002;
XX
XX 25-SEP-2001 (first entry)
XX
XX Human prostate-specific amino acid sequence L1-12/P501S.
XX
XX Human; prostate cancer; therapy; diagnosis; cat eye syndrome;
KW chromosome 22q11.2; prostate-specific protein; chromosome 1;
KW prostate specific antigen; PSA.
XX
XX Homo sapiens.
XX
XX WO200134802-A2.
XX
XX 17-MAY-2001.
XX
XX 09-NOV-2000; 2000WO-US30904.
XX
XX 12-NOV-1999; 99US-0439313.
XX
XX 18-NOV-1999; 99US-0443686.
XX
XX (CORI-) CORIXA CORP.
XX
XX Xu J, Dillon DC, Mitcham JL, Harlocker SL, Jiang Y, Reed SG;
PI Kalos MD, Retter MW, Stolk JA, Day CH, Skeiky YAW, Wang A;
XX
XX WPI: 2001-308785/32.
XX
XX Isolated polypeptide comprising at least an immunogenic portion of a

PT prostate-specific protein, useful in the diagnosis and therapy of
PT prostate cancer -
XX
XX Claim 3; Page 167-168; 325pp; English.
XX
XX The present invention describes an isolated polypeptide (P1) comprising
CC at least an immunogenic portion of a prostate-specific protein, or its
CC variant. Also described are polynucleotides (N1) encoding (P1), (P1) and
CC (N1) have cytostatic activity and can be used in vaccine production.
CC The polypeptides, nucleic acids and antibodies from the present
CC invention are useful in the diagnosis and therapy of prostate cancer.
CC Prostate specific genes P704P, P712P, P774P, P775P and B305D are located
CC in a genomic region on chromosome 22q11.2, known as the Cat Eye Syndrome
CC region. Prostate specific antigen (PSA) P501S was located on
CC chromosome 1. AAH84671 to AAH85143 and AAG99000 to AAG99077 represent
CC polynucleotide and polypeptide sequences used in the exemplification
CC of the present invention.
XX
XX Sequence 553 AA;

Query Match 53.6%; Score 15; DB 22; Length 553;
Best Local Similarity 100.0%; Pred. No. 8.8e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
|||||
Db 294 dfvgeglygvprae 308

RESULT 12
AAG62150
ID AAG62150 standard; Protein; 553 AA.
XX
XX AAG62150;
XX
XX 06-JUL-2001 (first entry)
XX
XX Human P501S inventive antigen SEQ ID NO: 333.
XX
XX Human; mouse; immunotherapy; cancer; leukaemia; WT1; Wilms' tumour gene;
KW chromosome 11p13; zinc finger transcription factor.
XX
XX Homo sapiens.
XX
XX WO200125273-A2.
XX
XX 12-APR-2001.
XX
XX 04-OCT-2000; 2000WO-US27465.
XX
XX 04-OCT-1999; 99US-0157459.
XX
XX (CORI-) CORIXA CORP.
XX
XX Skeiky YAW, Xu J, Cheever MA, Reed SG;
XX
XX WPI: 2001-328324/34.
XX
XX Polypeptide comprising part of the Wilms' Tumour gene product sequence is
PT used in the diagnosis and treatment of malignant diseases e.g. Leukemia
PT and cancer associated with WT1 -
XX
XX Disclosure; Page 212-213; 228pp; English.
XX
XX The present invention describes compositions comprising peptides derived
CC from the Wilms' tumour protein WT1 and methods for their use in treating
CC malignant diseases. Peptides derived from both the murine and human WT1
CC proteins are provided. The human WT1 gene is found on chromosome 11p13,
CC and the protein was shown to be a zinc finger transcription factor. The
CC immunogenic peptides of the invention are particularly useful in the
CC diagnosis and treatment of cancer and leukemia. The present sequence is
CC a polypeptide described in the exemplification of the invention.

DE Prostate tumour antigen amino acid sequence for a fusion protein.
 XX Human; prostate tumour antigen; prostate tumour; therapy; diagnosis;
 KM prostate cancer; immunogenic; cytostatic; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO200125272-A2.
 XX
 PD 12-APR-2001.
 XX
 PF 04-OCT-2000; 2000WO-US27464.
 XX
 PR 04-OCT-1999; 99US-0157455.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Xu J, Skeiky YAW, Reed SG, Cheever MA;
 XX
 DR WPI; 2001-245062/25.
 XX
 PT Prostate specific protein and its encoding polynucleotide, useful for
 XX the treatment and diagnosis of prostate cancer -
 PS Disclosure; Page 272-276; 276pp; English.
 XX
 CC The present invention describes an isolated polypeptide (I) comprising
 CC at least an immunogenic portion of a prostate tumour antigen protein or
 CC its variant. (I) have cytostatic activity and can be used in vaccine
 CC production. (I) prostate tumour antigen polynucleotides, an antigen
 CC presenting cell (APC e.g. a dendritic cell) that expresses (I), and a
 CC pharmaceutical composition containing (I) are useful for inhibiting the
 CC development of cancer in a patient. Antibodies specific for prostate
 CC specific proteins and oligonucleotides that hybridise to a
 CC polynucleotide that encodes a prostate specific protein are useful
 CC for detecting the presence or absence of a cancer or monitoring the
 CC progression the progression of a cancer, especially prostate cancer.
 CC AAH02422 to AAH2872, AAB74798 to AAB74821 and AAB74830 are sequences
 CC used in the exemplification of the present invention.
 CC
 XX
 SQ Sequence 1079 AA;

Query Match 53.6%; Score 15; DB 22; Length 1079;
 Best Local Similarity 100.0%; Pred. No. 1.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 DEVGEGLYGVPRAE 15
 ||||||||||||
 Db 820 divgeglygvprae 834

Search completed: June 26, 2002, 13:49:00
 Job time: 353 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:44:22 ; Search time 14.83 Seconds

(without alignments)
46.117 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28

Sequence: 1 DFVGEGLYGVPRAEGETEARHDEGVR 28

Scoring table: OLIGO

Searched: 231628 seqs, 24425594 residues

Word size : 0

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/2/1aa/PTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/2/1aa/Backfilest1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 15 | 53.6 | 553 | 4 | US-09-020-956-113 |
| 2 | 15 | 53.6 | 553 | 4 | US-09-030-607-113 |
| 3 | 15 | 53.6 | 553 | 4 | US-09-439-313-113 |
| 4 | 14 | 50.0 | 29 | 4 | US-09-439-313-546 |
| 5 | 13 | 46.4 | 25 | 4 | US-09-439-313-520 |
| 6 | 13 | 46.4 | 27 | 4 | US-09-071-710-39 |
| 7 | 13 | 46.4 | 27 | 4 | US-09-525-397-39 |
| 8 | 13 | 46.4 | 27 | 4 | US-09-439-313-566 |
| 9 | 13 | 46.4 | 255 | 4 | US-09-071-710-36 |
| 10 | 13 | 46.4 | 255 | 4 | US-09-525-397-36 |
| 11 | 11 | 39.3 | 15 | 4 | US-09-439-313-518 |
| 12 | 7 | 25.0 | 1938 | 1 | US-07-945-283-2 |
| 13 | 6 | 21.4 | 17 | 4 | US-09-439-313-519 |
| 14 | 6 | 21.4 | 327 | 4 | US-09-290-640-66 |
| 15 | 6 | 21.4 | 525 | 6 | 5183745-4 |
| 16 | 6 | 21.4 | 1489 | 6 | 5183745-2 |
| 17 | 6 | 21.4 | 1705 | 4 | US-08-668-785-4 |
| 18 | 6 | 21.4 | 1706 | 4 | US-08-668-785-2 |
| 19 | 6 | 21.4 | 1794 | 6 | 5183745-6 |
| 20 | 5 | 17.9 | 8 | 4 | US-09-420-211-12 |
| 21 | 5 | 17.9 | 9 | 2 | US-08-951-924A-3 |
| 22 | 5 | 17.9 | 11 | 3 | US-08-592-500-42 |
| 23 | 5 | 17.9 | 11 | 3 | US-08-195-006-42 |
| 24 | 5 | 17.9 | 11 | 3 | PCT-US94-07644A-42 |
| 25 | 5 | 17.9 | 14 | 1 | US-08-277-299-1 |
| 26 | 5 | 17.9 | 14 | 2 | US-08-880-671-1 |
| 27 | 5 | 17.9 | 14 | 4 | US-08-853-910-5 |

| | | | | | | |
|----|---|------|----|---|-------------------|--------------------|
| 28 | 5 | 17.9 | 15 | 1 | US-07-969-336-1 | Sequence 1, Appli |
| 29 | 5 | 17.9 | 15 | 2 | US-08-815-953-1 | Sequence 1, Appli |
| 30 | 5 | 17.9 | 17 | 2 | US-08-472-659-4 | Sequence 4, Appli |
| 31 | 5 | 17.9 | 17 | 2 | US-08-474-661-4 | Sequence 4, Appli |
| 32 | 5 | 17.9 | 17 | 2 | US-08-611-977-4 | Sequence 27, Appli |
| 33 | 5 | 17.9 | 18 | 1 | US-07-920-281C-27 | Sequence 27, Appli |
| 34 | 5 | 17.9 | 18 | 4 | US-08-466-277-27 | Sequence 8, Appli |
| 35 | 5 | 17.9 | 19 | 4 | US-08-975-040-8 | Sequence 1, Appli |
| 36 | 5 | 17.9 | 26 | 2 | US-08-563-892A-1 | Sequence 20, Appli |
| 37 | 5 | 17.9 | 26 | 2 | US-08-563-892A-20 | Sequence 21, Appli |
| 38 | 5 | 17.9 | 26 | 2 | US-08-563-892A-21 | Sequence 50, Appli |
| 39 | 5 | 17.9 | 26 | 2 | US-08-563-892A-22 | Sequence 50, Appli |
| 40 | 5 | 17.9 | 38 | 2 | US-08-284-391B-50 | Sequence 29, Appli |
| 41 | 5 | 17.9 | 38 | 4 | US-09-218-950-50 | Sequence 29, Appli |
| 42 | 5 | 17.9 | 53 | 2 | US-08-563-892A-29 | Sequence 14, Appli |
| 43 | 5 | 17.9 | 61 | 4 | US-09-537-357-13 | Sequence 27, Appli |
| 44 | 5 | 17.9 | 61 | 4 | US-09-537-357-14 | Sequence 27, Appli |
| 45 | 5 | 17.9 | 64 | 2 | US-08-563-892A-27 | Sequence 27, Appli |

ALIGNMENTS

```
RESULT 1
US-09-020-956-113
; Sequence 113, Application US/09020956
; Patent No. 6261562
;
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillin, Davin C.
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY OF PROSTATE CANCER AND METHODS
; NUMBER OF SEQUENCES: 178
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98104
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/020,956
; FILING DATE: 09-FEB-1998
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Makl, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.427C2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
;
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 553 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
;
; US-09-020-956-113
;
; Query Match 53.6%; Score 15; DB 4; Length 553;
; Best Local Similarity 100.0%; Pred. No. 2.6e-07;
; Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; 1 DFVGEGLYGVPRAE 15
; |||||||||||||
```

; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
 010131 43700
 010131 43700

;; CURRENT FILING DATE: 1999-11-12
;; NUMBER OF SEO ID NOS: 575

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 113
; ENCMNU, 553
```

```

; LENGTH: 003
; TYPE: PRT
;

```

```

; ORGANISM: Homo sapien
US-09-439-313-113

```

| | | | |
|-------------|---------|-----------|---------|
| Query Match | 53.6%; | Score 15; | DB 4; |
| Post Total | 100.0%; | Prod NO | 3 60-07 |

Best Local Similarity 100.00; Freq. NO. 2.00 07
Matches 15; Conservative 0; Mismatches 0

QY 1 DEVGEGLYQGVPAE 15

Dh 294 DEVGEGTYOGVPRAF 308

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

RESULT 4

US-09-439-313-546
; Sequence 546, Application US/09439313

Patent No. 6329505
GENERAL INFORMATION.

APPLICANT: Xu, Jiangchun

```

; APPLICANT: DILLON, DAVIN C.
; APPLICANT: MITCHAM, JENNIFER L.

```

```

; APPLICANT: Harlocker, Susan Louise
: APPLICANT: Jiaqg Yuquai

```

APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael

APPLICANT: Fanger, Gary

```

; APPLICANT: Ketter, Mark
; APPLICANT: Solk, John

```

APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR

TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121 427C9

CURRENT APPLICATION NUMBER: US/09/439,313

```

: CURRENT FILING DATE: 1999-11-12
:
: NUMBER OF SEQ ID NOS: 575

```

```
;; SOFTWARE: FastSeq for Windows Version 3.0
: SEO ID NO 546
```

```

LENGTH: 29
TYPE: PRT

```

ORGANISM: Homo sapiens

05-09-439-313-340

Query Match 50.0%; Score 14; DB 4;

```
Best Local Similarity 100.0%; Pred. No. 2.1e-07
Matches 14: Conservative 0: Mismatches 0
```

| ИДЕНТИФИКАЦИОННЫЙ № | ИДЕНТИФИКАЦИОННЫЙ № | ИДЕНТИФИКАЦИОННЫЙ № |
|---------------------|---------------------|---------------------|
| 471 | 471 | 471 |

QY 2 FVGGELIQQVPRAE 13
| | | | | | | | | |
| | | | | | | | | |

Db 1 FVGEGLYQGVPRAE 14

RESULTS

US-09-439-313-520

Sequence 520, Application US/094393A3
Patent No. 6329505

; GENERAL INFORMATION:
: APPLICANT: Xu, Jiangchun

APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.

APPLICANT: Harlocker, Susan Louise

```

; APPLICANT: JIANG YUQI
; APPLICANT: Reed, Steven G.

```

; APPLICANT: Kalos, Michael

APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solt, John
APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 520
LENGTH: 25
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-439-313-520

Query Match 46.4%; Score 13; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.7e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VEGGLYGVPRAE 15
|||||
DB 1 VEGGLYGVPRAE 13

RESULT 6
US-09-071-710-39
Sequence 39, Application US/09071710
Patent No. 6130043
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071.710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850,713
FILING DATE: 02-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:

TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6130043e
US-09-071-710-39

Query Match 46.4%; Score 13; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEARRHYDEGVR 28
|||||
DB 12 GTEARRHYDEGVR 24

RESULT 7
US-09-525-397-39
Sequence 39, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KLAAS, MICHAEL R.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
TITLE OF INVENTION: FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525.397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071.710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 27 amino acids
TYPE: amino acid

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-39

Query Match 46.4%; Score 13; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARRHYDEGVR 28
Db 12 GTEARRHYDEGVR 24

RESULT 8
US-09-439-313-566
Sequence 566, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, David C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqi
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Rafter, Mark
APPLICANT: Solik, John
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 566
LENGTH: 27
TYPE: PRT
ORGANISM: Homo sapiens
US-09-439-313-566

Query Match 46.4%; Score 13; DB 4; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.8e-06;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGEGLYQGVPRAE 15
Db 1 VGEGLYQGVPRAE 13

RESULT 9
US-09-071-710-36
Sequence 36, Application US/09071710
Patent No. 6130043
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE

NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA
ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,710
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/850,713
FILING DATE: 02-MAY-1997
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: NO. 6130043e
US-09-071-710-36

Query Match 46.4%; Score 13; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 GTEARRHYDEGVR 28
Db 12 GTEARRHYDEGVR 24

RESULT 10
US-09-525-397-36
Sequence 36, Application US/09525397
Patent No. 6252047
GENERAL INFORMATION:
APPLICANT: BILLING-MEDEL, PATRICIA
APPLICANT: COHEN, MAURICE
APPLICANT: COLPITTS, TRACEY L.
APPLICANT: FRIEDMAN, PAULA N.
APPLICANT: GORDON, JULIAN
APPLICANT: GRANADOS, EDWARD N.
APPLICANT: HODGES, STEVEN C.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: KRATOCHVIL, JON D.
APPLICANT: ROBERTS-RAPP, LISA
APPLICANT: RUSSELL, JOHN C.
APPLICANT: STROUPE, STEPHEN D.
TITLE OF INVENTION: REAGENTS AND METHODS USEFUL
FOR DETECTING DISEASES OF THE PROSTATE
NUMBER OF SEQUENCES: 41
CORRESPONDENCE ADDRESS:
ADDRESSEE: Abbott Laboratories
STREET: 100 Abbott Park Road
CITY: Abbott Park
STATE: IL
COUNTRY: USA

ZIP: 60064-3500
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/525,397
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/071,710
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Becker, Cheryl L.
REGISTRATION NUMBER: 35,441
REFERENCE/DOCKET NUMBER: 6083.US.P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 847/935-1729
TELEFAX: 847/938-2623
TELEX:
INFORMATION FOR SEQ ID NO: 36:
SEQUENCE CHARACTERISTICS:
LENGTH: 255 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: No. 6252047e
US-09-525-397-36

Query Match 46.4%; Score 13; DB 4; Length 255;
Best Local Similarity 100.0%; Pred. No. 1.2e-05;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 16 GTEARRHYDEGVR 28
Db 12 GTEARRHYDEGVR 24

RESULT 11
US-09-439-313-518
Sequence 518, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqun
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John
APPLICANT: Day, Crais
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 518
LENGTH: 15
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-439-313-518

Query Match 39.3%; Score 11; DB 4; Length 15;

Best Local Similarity 100.0%; Pred. No. 9.4e-05;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 16 GTEARRHYDEG 26
Db 5 GTEARRHYDEG 15

RESULT 12
US-07-945-283-2
Sequence 2, Application US/07945283
Patent No. 5352596
GENERAL INFORMATION:
APPLICANT: Cheung, Andrew K.
APPLICANT: Wesley, Ronald D.
TITLE OF INVENTION: Pseudotables Virus Deletion Mutants
TITLE OF INVENTION: Involving The EPO and LIT Genes
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Curtis P. Ribando
STREET: 1815 No. 5352596th University Street
CITY: Peoria
STATE: IL
COUNTRY: USA
ZIP: 61604
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/945,283
FILING DATE: 19920911
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Ribando, Curtis P.
REGISTRATION NUMBER: 27976
TELECOMMUNICATION INFORMATION:
TELEPHONE: 309-685-4128
TELEFAX: 309-685-4011 ext. 513
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1958 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-945-283-2

Query Match 25.0%; Score 7; DB 1; Length 1958;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 13 RAEGTEA 19
Db 537 RAEGTEA 543

RESULT 13
US-09-439-313-519
Sequence 519, Application US/09439313
Patent No. 6329505
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Dillon, Davin C.
APPLICANT: Mitcham, Jennifer L.
APPLICANT: Harlocker, Susan Louise
APPLICANT: Jiang Yuqun
APPLICANT: Reed, Steven G.
APPLICANT: Kalos, Michael
APPLICANT: Fanger, Gary
APPLICANT: Retter, Mark
APPLICANT: Solk, John

APPLICANT: Day, Craig
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
FILE REFERENCE: 210121.427C9
CURRENT APPLICATION NUMBER: US/09/439,313
CURRENT FILING DATE: 1999-11-12
NUMBER OF SEQ ID NOS: 575
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 519
LENGTH: 17
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Made in a lab
US-09-439-313-519

Query Match 21.4%: Score 6; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 16 GTEARR 21
|||
Db 5 GTEARR 10

RESULT 14
US-09-290-640-66
Sequence 66: Application US/09290640
Patent No. 6204055
GENERAL INFORMATION:
APPLICANT: Dean, Nicholas M.
APPLICANT: Marcussen, Eric G.
TITLE OF INVENTION: Antisense Compound Modulation of Fas Mediated Signaling
FILE REFERENCE: ISPH-0351
CURRENT APPLICATION NUMBER: US/09/290,640
CURRENT FILING DATE: 1999-04-12
NUMBER OF SEQ ID NOS: 85
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 66
LENGTH: 327
TYPE: PRT
ORGANISM: Mus musculus
US-09-290-640-66

Query Match 21.4%: Score 6; DB 4; Length 327;
Best Local Similarity 100.0%; Pred. No. 79;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EGLYOG 10
|||
Db 46 EGLYOG 51

RESULT 15
5183745-4
Patent No. 5183745
APPLICANT: DANCHIN, ANTOINE; GLASER, PHILIPPE; KRIN, EVELYN;
BARZU, OCTAVIEN; LADANT, DANIEL; ULLMAN, AGNES
TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
BIOLOGICAL USES
NUMBER OF SEQUENCES: 13
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,541
FILING DATE: 25-OCT-1989
SEQ ID NO: 4
LENGTH: 525
5183745-4

Query Match 21.4%: Score 6; DB 6; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 16 GTEARR 21
|||
Db 247 GTEARR 252

Search completed: June 26, 2002, 13:49:28
Job time: 306 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:45:47 ; Search time 20.39 seconds
(without alignments)
131.952 Million cell updates/sec

Title: US-09-838-785-25

Score: 28
Sequence: 1 DFVGEGLYQGVPRAGEFTEARRHYDEGVR 28

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 283138 seqs, 96089334 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|--------------------|
| 1 | 25.0 | 562 | 2 | H69545 | probable fatty-aci |
| 2 | 25.0 | 569 | 2 | C69471 | probable fatty-aci |
| 3 | 25.0 | 588 | 2 | A75282 | ABC transporter, A |
| 4 | 25.0 | 1473 | 1 | T31422 | C-terminal domain- |
| 5 | 25.0 | 1733 | 1 | B45344 | probable nuclear a |
| 6 | 25.0 | 1958 | 7 | B40505 | hypothetical prote |
| 7 | 25.0 | 1958 | 7 | B40505 | hypothetical prote |
| 8 | 21.4 | 777 | 2 | B82467 | hypothetical prote |
| 9 | 21.4 | 136 | 2 | A87681 | conserved hypotet |
| 10 | 21.4 | 219 | 2 | F75402 | probable competent |
| 11 | 21.4 | 231 | 2 | G87039 | conserved hypotet |
| 12 | 21.4 | 252 | 1 | A48008 | electron transfer |
| 13 | 21.4 | 264 | 2 | D97625 | hypothetical prote |
| 14 | 21.4 | 275 | 2 | H87538 | conserved hypotet |
| 15 | 21.4 | 293 | 2 | D89867 | hypothetical prote |
| 16 | 21.4 | 293 | 2 | A71946 | hypothetical prote |
| 17 | 21.4 | 298 | 2 | C64563 | conserved hypotet |
| 18 | 21.4 | 304 | 2 | H83636 | hypothetical prote |
| 19 | 21.4 | 306 | 2 | T26126 | hypothetical prote |
| 20 | 21.4 | 309 | 2 | AB1346 | hypothetical prote |
| 21 | 21.4 | 309 | 2 | AB1534 | oxidoreductases ho |
| 22 | 21.4 | 309 | 2 | AB1716 | oxidoreductases ho |
| 23 | 21.4 | 309 | 2 | AF1176 | hypothetical prote |
| 24 | 21.4 | 316 | 1 | D71246 | hypothetical prote |
| 25 | 21.4 | 316 | 2 | G72077 | hypothetical prote |
| 26 | 21.4 | 316 | 2 | G86345 | hypothetical prote |
| 27 | 21.4 | 318 | 2 | D75210 | hypothetical prote |
| 28 | 21.4 | 327 | 2 | A46484 | apoptosis-mediati |
| 29 | 21.4 | 348 | 2 | T03530 | cobM protein - Rho |

| | | | | | |
|----|------|-----|---|--------|---------------------|
| 30 | 21.4 | 352 | 2 | F64751 | probable ABC-type |
| 31 | 21.4 | 355 | 2 | AB3442 | response regulator |
| 32 | 21.4 | 358 | 2 | E95842 | probable lactose t |
| 33 | 21.4 | 395 | 2 | AE2143 | hypothetical prote |
| 34 | 21.4 | 418 | 2 | AD3417 | transporter, mfs s |
| 35 | 21.4 | 420 | 2 | AH1632 | isocitrate dehydrog |
| 36 | 21.4 | 458 | 2 | F75367 | potassium uptake p |
| 37 | 21.4 | 471 | 2 | T26612 | hypothetical prote |
| 38 | 21.4 | 483 | 2 | AB7583 | peptidoglycan bind |
| 39 | 21.4 | 488 | 2 | F86916 | NADH-dependent glu |
| 40 | 21.4 | 508 | 2 | AB4854 | hypothetical prote |
| 41 | 21.4 | 577 | 2 | H69354 | probable fatty-aci |
| 42 | 21.4 | 653 | 2 | D87602 | sensory box histid |
| 43 | 21.4 | 667 | 2 | AH2165 | bicarbonate transp |
| 44 | 21.4 | 686 | 2 | A34612 | zinc finger protei |
| 45 | 21.4 | 687 | 2 | B70515 | hypothetical prote |

ALIGNMENTS

RESULT 1
H69545
probable fatty-acid--CoA ligase (EC 6.2.1.-) fadD9 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: H69545
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
., J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: H69545
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-562 <KLE>
A:Cross-References: GB:AE001112; GB:AE000782; NID:92689435; PIDN:AA891290.1; PID:9265
C:Superfamily: 4-commarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase
F:87-553/Domain: acetate--CoA ligase homology <ACI>

Query Match 25.0%; Score 7; DB 2; Length 562;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 HYDEGVR 28
DB 27 HYDEGVR 33

RESULT 2
C69471
probable fatty-acid--CoA ligase (EC 6.2.1.-) fadD7 - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C:Accession: C69471
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
., J.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343
A:Accession: C69471
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-569 <KLE>
A:Cross-References: GB:AE000980; GB:AE000782; NID:92689303; PIDN:AA89478.1; PID:9264

C:Superfamily: 4-coumarate--CoA ligase; acetate--CoA ligase homology
C:Keywords: acid-thiol ligase
F:84-559/Domain: acetate--CoA ligase homology <ACLD>

Query Match 25.0%; Score 7; DB 2; Length 569;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 HYDEGVR 28
|||||
DB 24 HYDEGVR 30

RESULT 3

A75282

ABC transporter, ATP-binding protein, Msba family - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 26-May-2000

C:Accession: A75282

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uitterback, T.; Zalewski, C.; Ma

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896

A:Accession: A75282

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-588 <WHI>

A:Cross-references: GB:AE002066; GB:AE000513; NID:g6460186; PIDN:AEF11923.1; PID:g646019

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2379

A:Map position: 1

C:Superfamily: Escherichia coli ABC transporter mlaA; ATP-binding cassette homology

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 588;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 EGTEARR 21
|||||
DB 229 EGTEARR 235

RESULT 4

T31422

C-terminal domain-binding protein RA9 - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 07-Dec-1999

C:Accession: T31422

R:Yuryev, A.; Paturajan, M.; Littingtung, Y.; Joshi, R.V.; Gentile, C.; Gebara, M.; Cor

Proc. Natl. Acad. Sci. U.S.A. 93, 6975-6980, 1996

A:Title: The C-terminal domain of the largest subunit of RNA polymerase II interacts with

A:Reference number: 221024; MUID:96293459

A:Accession: T31422

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1473 <TDR>

A:Cross-references: EMBL:U49057; NID:g1438533; PID:g1438534; PIDN:AA052658.1

A:Experimental source: hippocampus

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 1473;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 10 GVPRAEG 16
|||||
DB 1276 GVPRAEG 1282

RESULT 5

B45344

probable nuclear antigen - suid herpesvirus 1 (strain Kaplan)

C:Species: suid herpesvirus 1

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Jul-1999

C:Accession: B45344

R:Vicek, C.; Kozmik, Z.; Paces, V.; Schirm, S.; Schwyzer, M.

Virology 179, 365-377, 1990

A:Title: Pseudorabies virus immediate-early gene overlaps with an oppositely oriented

A:Reference number: A45344; MUID:91021039

A:Accession: B45344

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-1733 <VIC>

A:Cross-references: GB:M34651; NID:g334070; PIDN:AAA47471.1; PID:g334072

C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match

Best Local Similarity 100.0%; Score 7; DB 1; Length 1733;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RAEGTEA 19
|||||
DB 326 RAEGTEA 332

RESULT 6

B40505

hypothetical protein - suid herpesvirus 1 (strain Indiana-Funkhouser or Becker)

C:Species: suid herpesvirus 1

C>Date: 10-Apr-1992 #sequence_revision 10-Apr-1992 #text_change 02-Sep-2000

C:Accession: B40505

R:Cheung, A.K.

J. Virol. 65, 5260-5271, 1991

A:Title: Cloning of the latency gene and the early protein 0 gene of pseudorabies vir

A:Reference number: A40505; MUID:91374576

A:Accession: B40505

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1958 <CHE>

A:Cross-references: GB:M57505; NID:g334066; PIDN:AAA47468.1; PID:g334068

C:Superfamily: pseudorabies virus 1 nuclear antigen

Query Match

Best Local Similarity 100.0%; Score 7; DB 2; Length 1958;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RAEGTEA 19
|||||
DB 537 RAEGTEA 543

RESULT 7

B82467

hypothetical protein VCA0389 [imported] - Vibrio cholerae (strain N16961 serogroup O1

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: B82467

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.

Charlson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406833

A:Accession: B82467

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-77 <HEI>

A:Cross-references: GB:AE004374; GB:AE003853; NID:g9657775; PIDN:AAF96295.1; GSPDB:GN

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:
A:Gene: VCA0389
A:Map position: 2

Query Match 21.4%; Score 6; DB 2; Length 77;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FYVEGL 7
|11111
DB 65 FYVEGL 70

RESULT 8
A87681
Conserved hypothetical protein CC3483 [Imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: A87681
R:Nierman, M.C.; Peltdlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwin, M.L.; Haft, D.H.; Kolon
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87681
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-136 <STO>
A:Cross-references: GB:AE005673; NID:g13425207; PIDN:AAK25445.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC3483

Query Match 21.4%; Score 6; DB 2; Length 136;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GLXGV 11
|11111
DB 65 GLXGV 70

RESULT 9
F75402
Probable competence protein ComF - *Deinococcus radiodurans* (strain R1)
C:Species: *Deinococcus radiodurans*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: F75402
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
A:Reference number: A75250; MUID:20036896
A:Accession: F75402
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-219 <WMT>
A:Cross-references: GB:AE001984; GB:AE000513; NID:g6459135; PIDN:AAF10957.1; PID:g645914
C:Genetics:
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1389
A:Map position: 1
C:Superfamily: transformation competence-related protein comF

Query Match 21.4%; Score 6; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 8 YGVPR 13

DB 71 YGVPR 76
|11111

RESULT 10
G87039
Conserved hypothetical protein ML1045 [Imported] - *Mycobacterium leprae*
C:Species: *Mycobacterium leprae*
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: G87039
R:Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.;
R.; Davies, R.M.; Devlin, K.; Duthey, S.; Fellwell, T.; Fraser, A.; Hamlin, N.; Holro
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.;
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: G87039
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: GB:AL450380; NID:g13093058; PIDN:CAC31426.1; GSPDB:GN00147
C:Genetics:
A:Gene: ML1045

Query Match 21.4%; Score 6; DB 2; Length 231;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 17 TEARRH 22
|11111
DB 203 TEARRH 208

RESULT 11
A48008
Electron transfer flavoprotein beta chain - *Paracoccus denitrificans*
C:Species: *Paracoccus denitrificans*
C>Date: 02-Jun-1995 #sequence_revision 12-Jul-1996 #text_change 08-May-1998
C:Accession: A48008; S23293
R:Bedzyk, L.A.; Escudero, K.W.; Gill, R.E.; Griffin, K.J.; Freeman, F.E.
J. Biol. Chem. 268, 20211-20217, 1993
A:Title: Cloning, sequencing, and expression of the genes encoding subunits of Paraco
A:Reference number: A48008; MUID:93388590
A:Accession: A48008
A:Molecule type: DNA
A:Residues: 1-252 <BED>
A:Cross-references: GB:U14864
R:Matmough, N.J.; Kiss, J.; Freeman, F.E.
Eur. J. Biochem. 205, 1089-1097, 1992
A:Title: Structural and redox relationships between *Paracoccus denitrificans*, porcine
A:Reference number: S23165; MUID:92249313
A:Accession: S23293
A:Molecule type: protein
A:Residues: 1-22 'R', 133-49 'AVA', 50-51 '74-82', 'Y', 164-167 'E', 169-183 '189-197' <WMT>
C:Complex: heterodimer of alpha (see PIR:B48008) and beta chains that binds one molec
C:Function:
A:Description: electron transfer from any of several dehydrogenases to the respirator
C:Superfamily: electron transfer flavoprotein beta chain
C:Keywords: electron transfer; FAD; flavoprotein; heterodimer
F.1-252/Product: electron transfer flavoprotein beta chain #status predicted <WMT>

Query Match 21.4%; Score 6; DB 1; Length 252;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RAEGTE 18
|11111
DB 110 RAEGTE 115

RESULT 12
D97625
hypothetical protein AGR_C.4025 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: D97625
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A:Reference number: A97359; PMID:11743194
A:Accession: D97625
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KOR>
A:Cross-references: GB:AE007869; PIDN:AAK87957.1; PID:g15157363; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C.4025
A:Map position: circular chromosome
C:Superfamily: streptomycetes coelicolor hypothetical protein SCA410.14c

Query Match 21.4%; Score 6; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EGLYOG 10
|||||
Db 36 EGLYOG 41

RESULT 13
AF2848
conserved hypothetical protein Atu2215 [imported] - Agrobacterium tumefaciens (strain C5
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AF2848
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCiell
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AF2848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-264 <KOR>
A:Cross-references: GB:AE008688; PIDN:AAU43204.1; PID:g17740685; GSPDB:GN00186
A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu2215
A:Map position: circular chromosome
C:Superfamily: streptomycetes coelicolor hypothetical protein SCA410.14c

Query Match 21.4%; Score 6; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EGLYOG 10
|||||
Db 36 EGLYOG 41

RESULT 14
H87538
hypothetical protein CC2337 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: H87538
R:Nierman, W.C.; Feidblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.

B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87538

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-275 <STO>
A:Cross-references: GB:AE005673; NID:g13423862; PIDN:AAK24308.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC2337
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 21.4%; Score 6; DB 2; Length 275;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 QGVPPA 14
|||||
Db 153 QGVPPA 158

RESULT 15
D89867
hypothetical protein SA0854 [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: D89867
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cut, L.; O
ma, A.; Mizutani-Oi, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: D89867
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-293 <KOR>
A:Cross-references: GB:BA000018; PID:g13700799; PIDN:BA842095.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: SA0854
C:Superfamily: oligopeptide permease protein oppB

Query Match 21.4%; Score 6; DB 2; Length 293;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FVGEGL 7
|||||
Db 278 FVGEGL 283

Search completed: June 26, 2002, 13:50:03
Job time: 256 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:48:22 : Search time 10.1 Seconds
(without alignments)
107.341 Million cell updates/sec

Title: US-09-838-785-25

Perfect score: 28

Sequence: 1 DFVGEGLYGVPRAGTEARHRYDEGVR 28

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 105224 seqs, 38719550 residues

Word size : 0

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|--------------|---------------------|
| 1 | 7 | 23.0 | 1733 | VNUA_PRVKA | P33485 pseudorabies |
| 2 | 6 | 21.4 | 252 | ETFB_PARDE | P38975 paracoccus |
| 3 | 6 | 21.4 | 293 | Y347_HELPY | Q92m1 helicobacte |
| 4 | 6 | 21.4 | 298 | Y347_HELPY | Q25114 helicobacte |
| 5 | 6 | 21.4 | 315 | MFT_HUMAN | Q9h2d1 homo sapien |
| 6 | 6 | 21.4 | 316 | Y441_CHLPP | Q928a2 chlamydia p |
| 7 | 6 | 21.4 | 327 | TNR6_MOUSE | P25446 mus musculu |
| 8 | 6 | 21.4 | 348 | AFUC_ECOLI | P37009 escherichia |
| 9 | 6 | 21.4 | 393 | ACK2_RHIME | Q9x449 rhizobium m |
| 10 | 6 | 21.4 | 686 | 1 ZN07_HUMAN | P17097 homo sapien |
| 11 | 6 | 21.4 | 840 | VPPI_YEAST | P32563 saccharomyc |
| 12 | 6 | 21.4 | 959 | SYL_AERPE | Q9y097 aeropyrum p |
| 13 | 6 | 21.4 | 1021 | CARA_MOUSE | P58660 mus musculu |
| 14 | 6 | 21.4 | 1363 | VGR3_MOUSE | P35917 mus musculu |
| 15 | 6 | 21.4 | 1705 | CYAA_BORBR | Q57506 bordetella |
| 16 | 6 | 21.4 | 1706 | CYAA_BORPE | P15318 bordetella |
| 17 | 5 | 17.9 | 60 | HMER_MYXCL | P31536 myxine glit |
| 18 | 5 | 17.9 | 64 | RL37_LYCES | P49212 lycopersico |
| 19 | 5 | 17.9 | 66 | ATPL_STRPN | Q59953 streptococc |
| 20 | 5 | 17.9 | 70 | CX2X_COMBE | Q9u323 conus betul |
| 21 | 5 | 17.9 | 71 | RECA_STRGC | P49986 streptococc |
| 22 | 5 | 17.9 | 72 | RL15_BACLI | P35138 bacillus 11 |
| 23 | 5 | 17.9 | 85 | COXG_BOVIN | P00429 bos taurus |
| 24 | 5 | 17.9 | 85 | COXG_HUMAN | P14854 homo sapien |
| 25 | 5 | 17.9 | 90 | PTSO_ECOLI | P33996 escherichia |
| 26 | 5 | 17.9 | 92 | DBR_CAUDR | O87475 caulobacter |
| 27 | 5 | 17.9 | 94 | YMFJ_ECOLI | P75973 escherichia |
| 28 | 5 | 17.9 | 102 | RECA_ENTFA | P42444 enterococcu |
| 29 | 5 | 17.9 | 103 | YH31_ARCFU | O28543 archaeoglob |
| 30 | 5 | 17.9 | 104 | RECA_LACDE | Q02347 lactobacilli |
| 31 | 5 | 17.9 | 104 | RECA_LACHE | Q02348 lactobacilli |
| 32 | 5 | 17.9 | 104 | RECA_LEUME | Q02349 leuconostoc |
| 33 | 5 | 17.9 | 104 | RECA_STRSL | P49987 streptococc |

| | | | | | | |
|----|---|------|-----|---|------------|--------------------|
| 34 | 5 | 17.9 | 106 | 1 | ATPR_DROME | Q24407 drosophila |
| 35 | 5 | 17.9 | 116 | 1 | Y998_RHIME | 052966 rhizobium m |
| 36 | 5 | 17.9 | 118 | 1 | VAMI_HUMAN | P23763 homo sapien |
| 37 | 5 | 17.9 | 118 | 1 | VAMI_MOUSE | Q62442 mus musculu |
| 38 | 5 | 17.9 | 118 | 1 | VAMI_RAT | Q63666 rattus norv |
| 39 | 5 | 17.9 | 118 | 1 | VPX_STVAI | Q02842 simian immu |
| 40 | 5 | 17.9 | 118 | 1 | Y357_VIBCH | Q9kx01 vibrio chol |
| 41 | 5 | 17.9 | 120 | 1 | CR13_ORYSA | Q06396 oryza sativ |
| 42 | 5 | 17.9 | 121 | 1 | RECA_RUMAL | P49985 ruminococcu |
| 43 | 5 | 17.9 | 126 | 1 | GP48_BPSPL | Q48402 bacterioph |
| 44 | 5 | 17.9 | 126 | 1 | YJ93_ARCFU | Q28286 archaeoglob |
| 45 | 5 | 17.9 | 131 | 1 | INL3_CALJA | Q97937 callithrix |

ALIGNMENTS

| RESULT | ID | STANDARD | PRT | Length |
|--|--|----------|-----|----------|
| 1 | VNUA_PRVKA | STANDARD | PRT | 1733 AA. |
| AC | P33485; | | | |
| DT | 01-FEB-1994 (Rel. 28, Created) | | | |
| DT | 01-FEB-1994 (Rel. 28, Last sequence update) | | | |
| DT | 01-FEB-1994 (Rel. 28, Last annotation update) | | | |
| DE | Probable nuclear antigen. | | | |
| OS | Pseudorabies virus (strain Kaplan) (PRV). | | | |
| OC | Viruses; dsDNA viruses, no RNA stage; Herpesviridae; | | | |
| OC | Alphaherpesvirinae; Varicellovirus. | | | |
| OX | NCBI_TaxID=33703; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE=91021039; PubMed=2171211; | | | |
| RA | Vicek C., Kozmik Z., Paces V., Schirm S., Schwytzer M.; | | | |
| RT | "Pseudorabies virus immediate-early gene overlaps with an oppositely | | | |
| RT | oriented open reading frame: characterization of their promoter and | | | |
| RT | enhancer regions."; | | | |
| RL | Virology 179:365-377(1990). | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-slb.ch/announce/ | | | |
| CC | or send an email to license@slb-slb.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; M34651; AAA47471.1; - | | | |
| DR | PIR; B45344; B45344. | | | |
| FT | DOMAIN 112 117 POLY-THR. | | | |
| FT | DOMAIN 179 1733 GLY-RICH. | | | |
| FT | DOMAIN 192 196 POLY-SER. | | | |
| FT | DOMAIN 271 298 POLY-PRO. | | | |
| FT | DOMAIN 304 308 POLY-ARG. | | | |
| FT | DOMAIN 883 889 POLY-GLY. | | | |
| FT | DOMAIN 1398 1405 POLY-GLY. | | | |
| SO | SEQUENCE 1733 AA; 172166 MW; 0C8CD8B5475BB5E2 CRC64; | | | |
| Query Match 25.0%; Score 7; DB 1; Length 1733; | | | | |
| Best Local Similarity 100.0%; Pred. No. 12; | | | | |
| Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0; | | | | |
| QY 13 RAEGTEA 19 | | | | |
| DB 326 RAEGTEA 332 | | | | |
| RESULT 2 | | | | |
| ETFB_PARDE STANDARD; PRT; 252 AA. | | | | |
| AC P38975; | | | | |
| DT 01-FEB-1995 (Rel. 31, Created) | | | | |
| DT 01-FEB-1995 (Rel. 31, Last sequence update) | | | | |

DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Electron transfer flavoprotein beta-subunit (beta-ETF) (Electron
 transfer flavoprotein small subunit) (ETFSS).
 GN ETPB.
 OS Paracoccus denitrificans.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
 OC Paracoccus.
 OX NCBI_TaxID=266;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-7.
 RC STRAIN-ATCC 13544;
 RX MEDLINE=93388590; PubMed=8376381;
 RA Bedzyk L.A., Escudero K.W., Gill R.E., Griffin K.J., Freeman F.E.;
 RT "Cloning, sequencing, and expression of the genes encoding subunits
 of Paracoccus denitrificans electron transfer flavoprotein.";
 RL J. Biol. Chem. 268:20211-20217(1993).
 RN [2]
 RP SEQUENCE OF 1-23; 33-51; 74-82; 163-183 AND 189-197.
 RC STRAIN-ATCC 13543;
 RX MEDLINE=92249313; PubMed=1576992;
 RA Watmough N.J., Kiss J., Freeman F.E.;
 RT "Structural and redox relationships between Paracoccus denitrificans,
 RT porcine and human electron-transferring flavoproteins.";
 RL Eur. J. Biochem. 205:1089-1097(1992).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
 RX MEDLINE=99151986; PubMed=10026281;
 RA Roberts D.L., Salazar D., Fulmer J.P., Freeman F.E., Kim J.-J.;
 RT "Crystal structure of Paracoccus denitrificans electron transfer
 RT flavoprotein: structural and electrostatic analysis of a conserved
 RT flavin binding domain.";
 RL Biochemistry 38:1977-1989(1999).
 CC -1- FUNCTION: THE ELECTRON TRANSFER FLAVOPROTEIN SERVES AS A SPECIFIC
 CC ELECTRON ACCEPTOR FOR OTHER DEHYDROGENASES. IT TRANSFERS THE
 CC ELECTRONS TO THE MAIN RESPIRATORY CHAIN VIA ETF-BIDIOINONE
 CC OXIDOREDUCTASE (ETF DEHYDROGENASE).
 CC -1- COPACITOR: CONTAINS ONE MOLECULE OF FAD PER DIMER.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE ETF BETA-SUBUNIT / FIXA FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L14864; AAA03071.1; -.
 DR PIR: S23293; S23293.
 DR PDB: IEP; 09-AUG-99.
 DR InterPro: IPR000049; ETF_beta.
 DR Pfam: PF01012; ETF_beta.1.
 DR ProDom: PD003528; ETF_beta.1.
 DR PROSITE: PS01065; ETF_beta.1.
 KW Electron transport; Flavoprotein; FAD; 3D-structure.
 FT CONFLICT 10 10 L -> V (IN REF. 2).
 FT CONFLICT 23 23 G -> R (IN REF. 2).
 FT CONFLICT 163 163 V -> Y (IN REF. 2).
 FT CONFLICT 168 168 Q -> E (IN REF. 2).
 SQ SEQUENCE 252 AA; 26673 MW; 00C43128BBA1DED CRC64;

Query Match 21.4%; Score 6; DB 1; Length 252;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 13 RAEGTE 18
 |||||
 DB 110 RAEGTE 115

RESULT 3

Y347_HELPJ
 ID Y347_HELPJ STANDARD: PRT; 293 AA.
 AC 09ZMA1.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein JHP0321.
 GN JHP0321.
 OS Helicobacter pylori J99 (Campylobacter pylori J99).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=85963;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99120557; PubMed=9923682;
 RA Alm R.A., Ling L.-S.L., Moir D.T., King B.L., Brown E.D., Dolg P.C.,
 RA Smith D.R., Noonan B., Guild B.C., deJonge B.L., Carmel G.,
 RA Tummlin P.J., Caruso A., Uria-Nickelsen M., Mills D.M., Ives C.,
 RA Gibson R., Merberg D., Mills S.D., Jiang O., Taylor D.E., Vovis G.F.,
 RA Trust T.J.;
 RT "Genomic sequence comparison of two unrelated isolates of the human
 RT gastric pathogen Helicobacter pylori.";
 RL Nature 397:176-180(1999).
 CC -1- SIMILARITY: BELONGS TO THE RUO FAMILY OF PSEUDOURIDINE SYNTHASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE001468; AAD05911.1; -.
 DR InterPro: IPR002990; PSI_RUO.
 DR InterPro: IPR000613; Pseudou_synth.
 DR Pfam: PF00849; Pseudou_synth_2; 1.
 DR ProDom: PD001819; Pseudou_synth; 1.
 DR PROSITE: PS01129; PSI_RUO; FALSE_NEG.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 293 AA; 33668 MW; D7EF3BC54A8E6168 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 293;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VEGELY 8
 |||||
 DB 234 VEGELY 239

RESULT 4
 Y347_HELPJ
 ID Y347_HELPJ STANDARD: PRT; 298 AA.
 AC 025114;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein HP0347.
 GN HP0347.
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=26695 / ATCC 700392;
 RX MEDLINE=97394467; PubMed=9252185;
 RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
 RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,
 RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
 RA Loftus B., Richardson D., Dodson R., Khairak H.G., Glodek A.,

RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
 RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
 RA Cotton M.D., Melman J.M., Fujii C., Bowman C., Matthey L., Wallin E.,
 RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
 RA Venter J.C.;
 RT "The complete genome sequence of the gastric pathogen *Helicobacter pylori*."
 RL Nature 388:539-547(1997).
 CC -1- SIMILARITY: BELONGS TO THE RLJ FAMILY OF PSEUDOURIDINE SYNTHASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE000551; AAD07404.1; -.
 DR TIGR: AP0347; -.
 DR InterPro: IPR002990; PSI_RLU.
 DR InterPro: IPR000613; Pseudou_synth.
 DR Pfam: PF00849; Pseudou_synth_2; 1.
 DR ProDom: PD001819; Pseudou_synth; 1.
 DR PROSITE: PS01129; PSI_RLU; FALSE_NEG.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 298 AA; 34347 MW; 116E4DD87779B15C CRC64;

Query Match 21.4%; Score 6; DB 1; Length 298;
 Best Local Similarity 100.0%; Pred. No. 27;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 VEGGLY 8
 |||||
 Db 234 VEGGLY 239

RESULT 5
 MFT_HUMAN STANDARD; PRT; 315 AA.
 AC Q9H2D1;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Mitochondrial folate transporter/carrier.
 GN MFT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20538421; PubMed=10978331;
 RA Titus S.A., Moran R.G.;
 RT "Retrovirally mediated complementation of the glyb phenotype. Cloning
 RT of a human gene encoding the carrier for entry of folates into
 RT mitochondria.";
 RL J. Biol. Chem. 275:36811-36817(2000).
 CC -1- FUNCTION: Transport folate across the inner membranes of
 CC mitochondria.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial
 CC inner membrane.
 CC -1- DOMAIN: COMPOSED OF THREE HOMOLOGOUS DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE MITOCHONDRIAL CARRIER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF283645; AAG37834.1; -.
 DR InterPro: IPR001993; Mitoch_carrier.
 DR InterPro: IPR002067; MLC_carrier.
 DR Pfam: PF00153; mito_carr; 3.
 DR PRINTS: PR00926; MITOCH_CARRIER.
 DR PROSITE: PS00215; MITOCH_CARRIER; 1.
 KW Mitochondrion; Inner membrane; Repeat; Transmembrane; Transport.
 FT TRANSMEM 89 106 POTENTIAL.
 FT TRANSMEM 227 243 POTENTIAL.
 FT TRANSMEM 281 300 POTENTIAL.
 SQ SEQUENCE 315 AA; 35388 MW; 60BEC0D61951EB6E CRC64;

Query Match 21.4%; Score 6; DB 1; Length 315;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GLYGV 11
 |||||
 Db 80 GLYGV 85

RESULT 6
 Y441_CHLPN STANDARD; PRT; 316 AA.
 ID Y441_CHLPN
 AC Q928A2; Q9JRW4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein CPN0441/CP0312/CPJ0441.
 GN CPN0441 OR CP0312 OR CPJ0441.
 OS Chlamydia pneumoniae (Chlamydia pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phyla.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CWL029;
 RX MEDLINE=99206606; PubMed=10192388;
 RA Kalman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W.,
 RA Olinger L., Grimwood J., Davis R.W., Stephens R.S.;
 RT "Comparative genomes of Chlamydia pneumoniae and C. trachomatis.";
 RL Nat. Genet. 21:385-389(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S.,
 RA Liner K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwinn M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10871362;
 RA Shirai M., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kohara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA.";
 RL Nucleic Acids Res. 28:2311-2314(2000).
 CC -1- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0441/CT007/TC0275
 CC FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AE001627; AAD18585.1; -
 DR EMBL: AE002194; AAF73653.1; -
 DR EMBL: AP002546; BAA98649.1; -
 DR TIGR: CP0312; -
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 243 243 Y->H (IN REF. 1).
 SO SEQUENCE 316 AA; 35415 MW; C821CEB53B014B41 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 316;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GLYOGV 11
 DB 160 GLYOGV 165

RESULT 7
 TNR6_MOUSE STANDARD; PRT; 327 AA.

AC P25446;
 DT 01-MAY-1992 (Rel. 22, Created)
 DT 01-MAY-1992 (Rel. 22, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Tumor necrosis factor receptor superfamily member 6 precursor (FASL
 DE receptor) (Apoptosis-modulating surface antigen FAS) (Apo-1 antigen)
 DE (CD95).
 GN TNFRSF6 OR APT1 OR FAS.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 CX [1]
 RX SEQUENCE FROM N.A.
 RA MEDLINE=92148151; PubMed=1371136;
 RA Watanabe-Fukunaga R., Brannan C.I., Itoh N., Yonehara S.,
 RA Watanabe N.G., Jenkins N.A., Nagata S.,
 RA "The CDNA structure, expression, and chromosomal assignment of the
 RA mouse Fas antigen.";
 RT J. Immunol. 148:1274-1279(1992).
 RL [2]
 RP SEQUENCE OF 1-96 FROM N.A.
 RX MEDLINE=93189576; PubMed=7680478;
 RA Adachi M., Watanabe-Fukunaga R., Nagata S.;
 RA "Aberrant transcription caused by the insertion of an early
 RA transposable element in an intron of the Fas antigen gene of 1pr
 RA mice.";
 RT Proc. Natl. Acad. Sci. U.S.A. 90:1756-1760(1993).
 RL [3]
 RP VARIANT LPR.
 RX MEDLINE=92195401; PubMed=1372394;
 RA Watanabe-Fukunaga R., Brannan C.I., Copeland N.G., Jenkins N.A.,
 RA Nagata S.;
 RA "Lymphoproliferation disorder in mice explained by defects in Fas
 RA antigen that mediates apoptosis.";
 RT Nature 356:314-317(1992).
 RL [4]
 CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS FASL. THE
 CC ADAPTOR MOLECULE FADD RECRUITS CASPASE-8 TO THE ACTIVATED
 CC RECEPTOR. THE RESULTING AGGREGATE CALLED THE DEATH-INDUCING
 CC SIGNALING COMPLEX (DISC) PERFORMS CASPASE-8 PROTEOLYTIC
 CC ACTIVATION. ACTIVE CASPASE-8 INITIATES THE SUBSEQUENT CASCADE OF
 CC CASPASES (ASPARTATE-SPECIFIC CYSTEINE PROTEASES) MEDIATING
 CC APOPTOSIS. FAS-MEDIATED APOPTOSIS MAY HAVE A ROLE IN THE
 CC INDUCTION OF PERIPHERAL TOLERANCE, IN THE ANTIGEN-STIMULATED
 CC SUICIDE OF MATURE T-CELLS, OR BOTH (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: DETECTED IN VARIOUS TISSUES INCLUDING THYMUS,
 CC LIVER, LUNG, HEART, AND ADULT OVARY.
 CC -1- DOMAIN: CONTAINS A DEATH DOMAIN INVOLVED IN THE BINDING OF FADD,

CC AND MAYBE TO OTHER CYTOSOLIC ADAPTOR PROTEINS.
 CC -1- DISEASE: DEFECTS IN TNFRSF6 ARE THE CAUSE OF A LYMPHOPROLIFERATION
 CC DISORDER (LPR) RESPONSIBLE FOR LYMPHADENOPATHY AND AUTOANTIBODY
 CC PRODUCTION.
 CC -1- SIMILARITY: CONTAINS 3 TNFR-CYS REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 DEATH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M83649; AAA37593.1; -
 DR EMBL: S56490; AAB25700.1; -
 DR EMBL: S56485; AAB25700.1; JOINED.
 DR EMBL: S56486; AAB25700.1; JOINED.
 DR PIR: A46484; A46484.
 DR HSSP: P25445; 1DDF.
 DR MGD: MGI:95484; Tnfrsf6.
 DR InterPro: IPR000488; Death.
 DR InterPro: IPR001368; TNFR_c6.
 DR Pfam: PF00531; death; 1.
 DR Pfam: PF00020; TNFR_c6; 3.
 DR SMART: SM00005; DEATH; 1.
 DR SMART: SM00208; TNFR; 3.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS50050; TNFR_NGFR_2; 2.
 DR PROSITE: PS50017; DEATH_DOMAIN; 1.
 KW Apoptosis; Receptor; Glycoprotein; Transmembrane; Repeat; Signal;
 KW Disease mutation.
 FT SIGNAL 1 21
 FT CHAIN 22 327
 FT FT
 FT DOMAIN 22 169 TUMOR NECROSIS FACTOR RECEPTOR
 FT TRANSMEM 170 186 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 187 327 POTENTIAL.
 FT REPEAT 43 79 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 80 123 TNFR-CYS 1.
 FT REPEAT 124 162 TNFR-CYS 2.
 FT REPEAT 222 306 TNFR-CYS 3.
 FT DOMAIN 222 306 DEATH.
 FT CARBOHYD 43 43 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 246 246 I -> N (IN LPR).
 SO SEQUENCE 327 AA; 37418 MW; F6BFFC5ACE356EEB CRC64;

Query Match 21.4%; Score 6; DB 1; Length 327;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 5 EGLYOG 10
 DB 46 EGLYOG 51

RESULT 8
 AFUC_ECOLI STANDARD; PRT; 348 AA.

ID AFUC_ECOLI
 AC P37009; P77157;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative ferric transport ATP-binding protein afuc.
 DE ASFC OR B0262.
 GN ASFC
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 CC NCB1_TaxID=562;
 CX [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-K12 / MG1655;
 RX MEDLINE-97426617; PubMed-9278503;
 RA Blatner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of *Escherichia coli* K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12 / W3110;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
 RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
 RT "Systematic sequencing of the *Escherichia coli* genome: analysis of the
 RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Lasker D., Lew H., Lin D., Namath A., Oetner P., Roberts D.,
 RA Davis R.W.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-220 FROM N.A.
 RX MEDLINE-94156632; PubMed-8113168;
 RA Volker M.R., Loewen P.C., Swatara J., Crowley D., Conley M.;
 RT "The delta (argf-lacZ)205(u169) deletion greatly enhances resistance
 RT to hydrogen peroxide in stationary-phase *Escherichia coli*.";
 RL J. Bacteriol. 176:1297-1302(1994).
 CC -1- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
 CC FOR IRON. PROBABLY RESPONSIBLE FOR ENERGY COUPLING TO THE
 CC TRANSPORT SYSTEM.
 CC -1- MISCELLANEOUS: IN E. COLI, THE AFU SYSTEM SEEMS TO HAVE BEEN
 CC DELETED BY AN INSERTION SEQUENCE. AFUA IS TOTALLY LOST, AFUB IS
 CC PARTIALLY PRESENT AND AFUC IS TOTALLY CONSERVED.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -1- CAUTION: REF.4 DIFFERS FROM THAT SHOWN FROM POSITION 147 TO 220
 CC DUE TO A FRAMESHIFT.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE000134; AAC73365.1; ALT_INIT.
 DR EMBL; D83536; BAA77930.1; -;
 DR EMBL; U70214; AAB08683.1; ALT_INIT.
 DR EMBL; L20943; -; NOT_ANNOTATED_CDS.
 DR Ecocore; EG12340; afuc.
 DR InterPro; IPR003593; AAA.
 DR InterPro; IPR003439; ABC_transportr.
 DR InterPro; IPR001687; ATP_GTP_A.
 DR Pfam; PF00005; ABC_tran; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER; 1.
 KW Iron transport; Transport; ATP-binding; Complete proteome.
 FT NP_BIND 39 46 ATP (POTENTIAL).
 FT CONFLICT 147 148 LA -> WP (IN REF. 2).
 FT SEQUENCE 348 AA; 39059 MW; DAD33754217DAIA CRC64;

Query Match 21.4%; Score 6; DB 1; Length 348;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 10 GVPRAE 15
 Db 107 GVPRAE 112

RESULT 9
 ACK2_RHIME STANDARD; PRT; 393 AA.
 ID ACK2_RHIME
 AC 09X449; 09XDG2;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Acetate kinase (EC 2.7.2.1) (Acetokinese).
 GN ACKA.
 OS Rhizobium meliloti (Sinorhizobium meliloti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-104A14;
 RX MEDLINE-99194732; PubMed-10094701;
 RA Summers M.L., Denton M.C., McDermott T.R.;
 RT "Genes coding for phosphotransacetylase and acetate kinase in
 RT Sinorhizobium meliloti are in an operon that is inducible by phosphate
 RT stress and controlled by phoB.";
 RL J. Bacteriol. 181:2217-2224(1999).
 RN [2]
 RP SEQUENCE OF 207-314 FROM N.A.
 RC STRAIN-104A14;
 RX MEDLINE-99022212; PubMed-9805396;
 RA Summers M.L., Elkins J.G., Elliot B.A., McDermott T.R.;
 RT "Expression and regulation of phosphate stress inducible genes in
 RT Sinorhizobium meliloti.";
 RL Mol. Plant Microbe Interact. 11:1094-1101(1998).
 CC -1- CATALYTIC ACTIVITY: ATP + acetate -> ADP + acetyl phosphate.
 CC -1- PATHWAY: FIRST OF TWO STEPS IN THE CONVERSION OF ACETATE TO
 CC ACETYL-COA.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE ACETOKINASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF095903; AAD24358.1; -;
 DR EMBL; AF074452; AAD42996.1; -;
 DR InterPro; IPR000890; Acetate_kin.
 DR Pfam; PF00871; Acetate_kinase; 1.
 DR PRINTS; PRO0471; ACETATEKINASE.
 DR PROSITE; PS01075; ACETATE_KINASE_1; 1.
 DR PROSITE; PS01076; ACETATE_KINASE_2; FALSE_NEG.
 KW Transferase; Kinase.
 FT CONFLICT 209 212 SCAS -> AEFR (IN REF. 2).
 FT CONFLICT 227 238 LHRPSTGCMWD -> GFTALDGLPMGTR (IN REF.
 FT SEQUENCE 393 AA; 42135 MW; 48FD185524CDD64C CRC64;

Query Match 21.4%; Score 6; DB 1; Length 393;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 23 YDEGVR 28
 Db 169 YDEGVR 174

RESULT 10
 ZN07_HUMAN STANDARD; PRT; 686 AA.
 ID ZN07_HUMAN
 AC P17097; P17015;

DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Zinc finger protein 7 (Zinc finger protein KOX4) (Zinc finger protein
 HE 16).
 GN ZNF7 OR KOX4.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=90169993; PubMed=2106481;
 RA Lania L., Donti E., Pannuti A., Pascucci A., Penque G.,
 RT Felicelli I., La Manta G., Lanfranco L., Pelicci P.-G.;
 RT "cDNA isolation, expression analysis, and chromosomal localization of
 RT two human zinc finger genes.";
 RL Genomics 6:333-340(1990).
 RN [2]
 RP SEQUENCE OF 413-468 FROM N.A.
 RC TISSUE=Lymphoid;
 RX MEDLINE=9145339; PubMed=2288909;
 RA Thiesen H.-J.;
 RT "Multiple genes encoding zinc finger domains are expressed in human T
 RT cells.";
 RL New Biol. 2:363-374(1990).
 CC -1- FUNCTION: MAY FUNCTION AS A TRANSCRIPTION FACTOR.
 CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).
 CC -1- TISSUE SPECIFICITY: UBIQUITOUSLY PRESENT IN MANY HUMAN CELL LINES
 CC -1- OF DIFFERENT EMBRYOLOGICAL DERIVATION.
 CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
 CC FINGER PROTEINS.
 CC -1- SIMILARITY: CONTAINS 1 KRAB DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M29580; AAA61313.1; -;
 DR EMBL; X52335; CAA36561.1; -;
 DR PIR; A34612; A34612.
 DR PIR; S10421; S10421.
 DR HSSP; P08047; 1SP2.
 DR MIM; 194531; -;
 DR InterPro: IPR001909; KRAB.
 DR InterPro: IPR000822; Znf-C2H2.
 DR Pfam; PF01352; KRAB; 1.
 DR Pfam; PF00096; Zf-C2H2; 14.
 DR PRINTS; PR00048; ZINCFINGER.
 DR SMART; SM00349; KRAB; 1.
 DR SMART; SM00355; Znf-C2H2; 14.
 DR PROSITE; PS50805; KRAB; 1.
 DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 14.
 DR PROSITE; PS50157; ZINC_FINGER_C2H2_2; 14.
 KW Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 KW Nuclear protein; Repeat.
 FT DOMAIN 4 76 KRAB.
 FT ZN_FING 223 684 ZINC_FINGERS.
 FT ZN_FING 223 245 C2H2-TYPE.
 FT ZN_FING 250 272 C2H2-TYPE.
 FT ZN_FING 278 300 C2H2-TYPE.
 FT ZN_FING 306 328 C2H2-TYPE.
 FT ZN_FING 328 356 C2H2-TYPE.
 FT ZN_FING 356 384 C2H2-TYPE.
 FT ZN_FING 384 435 C2H2-TYPE.
 FT ZN_FING 435 463 C2H2-TYPE.
 FT ZN_FING 463 491 C2H2-TYPE.
 FT ZN_FING 497 519 C2H2-TYPE.

FT ZN_FING 525 547 C2H2-TYPE.
 FT ZN_FING 553 575 C2H2-TYPE.
 FT ZN_FING 581 603 C2H2-TYPE.
 FT ZN_FING 634 656 C2H2-TYPE.
 FT ZN_FING 662 684 C2H2-TYPE.
 SO SEQUENCE 686 AA; 77887 MW; 0397ADBAEFDF4BFC CRC64;

 Query Match 21.4%; Score 6; DB 1; Length 686;
 Best Local Similarity 100.0%; Pred. No. 56;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 14 AEGTEA 19
 DB 71 AEGTEA 76

 RESULT 11
 VPH1_YEAST STANDARD; PRT; 840 AA.
 ID VPH1_YEAST
 AC P32563;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-OCT-1993 (Rel. 27, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Vacuolar ATP synthase 95 kDa subunit (vacuolar ATPase 95 kDa subunit).
 GN VPH1 OR YOR270C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 CC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 CC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 CX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92332542; PubMed=1385813;
 RA Manolson M.F., Proteau D., Preston R.A., Stenbit A., Roberts B.T.,
 RA Hoyt M.A., Preuss D., Mulholland J., Botstein D., Jones E.W.;
 RT "The VPH1 gene encodes a 95-kDa integral membrane polypeptide
 RT required for in vivo assembly and activity of the yeast vacuolar
 RT H(+)-ATPase.";
 RL J. Biol. Chem. 267:14294-14303(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93147685; PubMed=1491220;
 RA Manolson M.F., Proteau D., Jones E.W.;
 RT "Evidence for a conserved 95-120 kDa subunit associated with and
 RT essential for activity of V-ATPases.";
 RL J. Exp. Biol. 172:105-112(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=97051594; PubMed=8896271;
 RA Cheret G., Bernardi A., Sor F.J.;
 RT "DNA sequence analysis of the VPH1-SNF2 region on chromosome XV of
 RT Saccharomyces cerevisiae.";
 RL Yeast 12:1059-1064(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C / FY1679;
 RX MEDLINE=97298311; PubMed=9153759;
 RA Poljey R., Janniaux J.C.;
 RT "Sequencing analysis of a 36.8 kb fragment of yeast chromosome XV
 RT reveals 26 open reading frames including SEC63, CDC31, SUG2, GCD1,
 RT Yea2 13:483-487(1997).
 RL Yeast 13:483-487(1997).
 CC -1- FUNCTION: REQUIRED FOR ASSEMBLY AND ACTIVITY OF THE VACUOLAR
 CC ATPASE. POTENTIAL ROLE IN DIFFERENTIAL TARGETING AND REGULATION OF
 CC THE ENZYME FOR A SPECIFIC ORGANELLE.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
 CC -1- SIMILARITY: BELONGS TO THE VAPPAE 116 kDa SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: M89778; AAA35211.1; -
 DR EMBL: X89633; CAA61776.1; -
 DR EMBL: 275178; CAA99494.1; -
 DR EMBL: 275179; CAA99496.1; -
 DR PIR: A42970; A42970.
 DR SGD: S0005796; VPH1.
 DR InterPro: IPR002490; V_ATPase_sub.a.
 DR Pfam: PF01496; V_ATPase_sub.a; 1.
 KW Hydrogen ion transport; Transmembrane; Glycoprotein.
 FT DOMAIN 1 411
 FT TRANSMEM 412 432
 FT DOMAIN 433 462
 FT TRANSMEM 463 483
 FT DOMAIN 484 540
 FT TRANSMEM 541 561
 FT DOMAIN 562 571
 FT TRANSMEM 572 592
 FT DOMAIN 593 635
 FT TRANSMEM 636 656
 FT DOMAIN 657 760
 FT TRANSMEM 761 787
 FT DOMAIN 788 840
 FT CARBOHYD 113 113
 FT CARBOHYD 280
 FT CARBOHYD 324
 SQ SEQUENCE 840 AA; 95528 MW; 77709A914410CD4D CRC64;

Query Match 21.4%; Score 6; DB 1; Length 840;
 Best Local Similarity 100.0%; Pred. No. 67;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 FVGEGL 7
 Db 810 FVGEGL 815

RESULT 12
 SYL_AERPE STANDARD; PRT; 959 AA.
 AC Q9YD97;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Leucyl-tRNA synthetase (EC 6.1.1.4) (Leucine--tRNA ligase) (LeuRS).
 GN LEUS OR APL1015.
 OS Aeropyrum pernix.
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;
 OC Aeropyrum.
 OX NCBI_TaxID=56636;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1;
 RX MEDLINE=99310339; PubMed=10382966;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
 RA Jin-ou K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,
 RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
 RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
 RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
 RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.,
 RT "Complete genome sequence of an aerobic hyper-thermophilic
 RT crenarchaeon, Aeropyrum pernix K1.";
 RL DNA Res. 6:83-101(1999).
 CC -1- CATALYTIC ACTIVITY: ATP + L-leucine + tRNA(Leu) = AMP +
 CC diphosphate + L-leucyl-tRNA(Leu).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: APO0060; BAA80000.1; -
 DR InterPro: IPR002300; tRNA-synt_1a.
 DR InterPro: IPR001412; tRNA-synt_1.
 DR Pfam: PF00133; tRNA-synt_1; 1.
 DR PROSITE: PS00178; AA_tRNA_LIGASE_1; FALSE_NEG.
 KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
 FT SITE 39 49
 FT SITE 637 641
 FT BINDING 640 640
 SQ SEQUENCE 959 AA; 110519 MW; 985CB5168728F7AB CRC64;

Query Match 21.4%; Score 6; DB 1; Length 959;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 11 VPRAEG 16
 Db 357 VPRAEG 362

RESULT 13
 CARA_MOUSE STANDARD; PRT; 1021 AA.
 AC P58660;
 DT 01-MAR-2002 (Rel. 41, Created)
 DT 01-MAR-2002 (Rel. 41, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Caspase recruitment domain protein 10 (Bcl10-interacting MAGUK protein
 DE 1) (Bim1).
 GN CARD10 OR BIM1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21391892; PubMed=11387339;
 RA McAllister-Lucas L.M., Inohara N., Lucas P.C., Ruland J., Benito A.,
 RA Li Q., Chen S., Chen F.F., Yamaoka S., Verma I.M., Mak T.W.,
 RA Nunez G.;
 RT "Bim1, a MAGUK family member linking protein kinase c activation to
 RT Bcl10-mediated NF-kappa B induction.";
 RL J. Biol. Chem. 276:30589-30597(2001).
 CC -1- FUNCTION: Activates NF-kappaB via Bcl10 and IKK.
 CC -1- SUBUNIT: CARD10 and Bcl10 bind to each other by CARD-CARD
 CC interaction. They both participate in a complex with MAL1, where
 CC MAL1 binds to Bcl10.
 CC -1- TISSUE SPECIFICITY: Highly expressed in kidney, heart followed by
 CC brain, lung, liver, skeletal muscle and testis.
 CC -1- SIMILARITY: CONTAINS 1 CARD DOMAIN.
 CC -1- CAUTION: Supposed to contain a SH3, a PDZ and a guanylate kinase-
 CC like domain. But none of these 3 domains are detected by PROSITE,
 CC Pfam or SMART.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF363456; AAK60136.1; -
 DR PROSITE: PS50209; CARD; 1.

KW Coiled coil. 23 115 CAND.
 FT DOMAIN 138 450 COILED COIL (POTENTIAL).
 FT DOMAIN 558 565 POLY-SER.
 SQ SEQUENCE 1021 AA; 114413 MW; 4811A09BDB8F792C CRC64;

Query Match 21.4%; Score 6; DB 1; Length 1021;
 Best Local Similarity 100.0%; Pred. No. 80;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 12 PRACT 17
 DB 641 PRACT 646

RESULT 14
 VGR3_MOUSE STANDARD; PRT; 1363 AA.
 ID VGR3_MOUSE
 AC P35917;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 01-MAR-2002 (Rel. 41, Last annotation update)
 DE Vascular endothelial growth factor receptor 3 precursor (EC 2.7.1.112)
 DE (VEGF-3) (tyrosine-protein kinase receptor FLT4).
 GN FLT4 OR FLT-4.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=93330572; PubMed=8393164;
 RA Flinerty H., Kelleher K., Morris G.E., Bean K., Merberg D.M.,
 RA Kriz R., Morris J.C., Sockdeo H., Turner K.J., Wood C.R.;
 RT Molecular cloning of murine FLT and FLT4.*;
 RL Oncogene 8:2293-2298(1993).
 CC -1- FUNCTION: RECEPTOR FOR VEGF-C. HAS A TYROSINE-PROTEIN KINASE
 CC ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT LUNG AND LIVER, AND IN
 CC FETAL LIVER, BRAIN, INTERESTINE AND PLACENTA.
 CC -1- SIMILARITY: BELONGS TO THE CSF-1/PDGF RECEPTOR FAMILY OF TYROSINE-
 CC PROTEIN KINASES.
 CC -1- SIMILARITY: CONTAINS 7 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L07296; AAA40077.1; -;
 DR HSSP: P11362; 1FGK.
 DR MCD: MGI:95561; P1f4.
 DR InterPro: IPR000719; Euk_Pkinase.
 DR InterPro: IPR003506; Ig_MHC.
 DR InterPro: IPR003598; Ig_C2.
 DR InterPro: IPR003600; Ig_Like.
 DR InterPro: IPR01824; Receptor_tyr_kin_III.
 DR InterPro: IPR01245; Tyr_Pkinase.
 DR Pfam: PF00047; Ig; 6.
 DR Pfam: PF00069; Pkinase; 2.
 DR SMART: SM00410; IG_Like; 3.
 DR SMART: SM00408; ICG2; 2.
 DR SMART: SM00219; TyrKc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE: PS00240; RECEPTOR_TYR_KIN_III; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Tyrosine-protein kinase; Phosphorylation; ATP-binding;
 KW Receptor; Transmembrane; Signal; Immunoglobulin domain; Repeat;
 KW Glycoprotein.
 FT SIGNAL 1 24
 FT CHAIN 25 1363
 FT DOMAIN 25 775
 FT TRANSMEM 776 797
 FT DOMAIN 798 1363
 FT DOMAIN 44 118
 FT DOMAIN 151 213
 FT DOMAIN 245 317
 FT DOMAIN 351 403
 FT DOMAIN 438 541
 FT DOMAIN 571 660
 FT DOMAIN 692 758
 FT DOMAIN 845 1173
 FT NP_BIND 851 859
 FT BINDING 879 879
 FT ACT_SITE 1037 1037
 FT DISULFID 51 111
 FT DISULFID 158 206
 FT DISULFID 252 310
 FT DISULFID 445 534
 FT DISULFID 578 653
 FT DISULFID 699 751
 FT MOD_RES 1068 1068
 FT CARBOHYD 33 33
 FT CARBOHYD 104 104
 FT CARBOHYD 166 166
 FT CARBOHYD 251 251
 FT CARBOHYD 299 299
 FT CARBOHYD 411 411
 FT CARBOHYD 515 515
 FT CARBOHYD 527 527
 FT CARBOHYD 582 582
 FT CARBOHYD 594 594
 FT CARBOHYD 683 683
 FT CARBOHYD 690 690
 FT CARBOHYD 758 758
 SQ SEQUENCE 1363 AA; 153015 MW; F1BF8A2BDBF99BE9 CRC64;

Query Match 21.4%; Score 6; DB 1; Length 1363;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 ECTEAR 20
 DB 85 ECTEAR 90

RESULT 15
 CYAA_BORBR STANDARD; PRT; 1705 AA.
 ID CYAA_BORBR
 AC Q57506; O05179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Bifunctional hemolysin-adenylate cyclase precursor (Cyclolysin) (ACT)
 DE (AC-HLY) (Contains: Calmodulin-sensitive adenylylate cyclase
 DE (EC 4.6.1.1) (ATP pyrophosphatase) (Adenylyl cyclase); Hemolysin).
 GN CYA OR CYAA.
 OS Bacteriella bronchiseptica (Alcaligenes bronchisepticus).
 CC Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae;
 CC Bordetella.
 OX NCBI_TaxID=518;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CIP 9.73;
 RX MEDLINE=96009899; PubMed=7557410;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 26, 2002, 13:47:32 ; Search time 26.53 Seconds
(without alignments)
182.580 Million cell updates/sec

Title: US-09-838-785-25
Perfect score: 28
Sequence: 1 DFVGEGLYOGVPRAEGRHRYDEGVR 28

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 562222 seqs, 172994929 residues

Word size : 0

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL_19:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeal:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------|
| 1 | 15 | 53.6 | 501 | 6 | 095KC5 macaca fasc |
| 2 | 15 | 53.6 | 553 | 4 | 096JT2 |
| 3 | 15 | 53.6 | 553 | 6 | 095K15 macaca fasc |
| 4 | 7 | 25.0 | 179 | 5 | 09VSG8 |
| 5 | 7 | 25.0 | 364 | 2 | 09RJ80 |
| 6 | 7 | 25.0 | 562 | 17 | 0930302 |
| 7 | 7 | 25.0 | 569 | 17 | 028502 |
| 8 | 7 | 25.0 | 588 | 16 | 09RNV6 |
| 9 | 7 | 25.0 | 1473 | 11 | 063625 |
| 10 | 7 | 25.0 | 1958 | 12 | 069340 |
| 11 | 6 | 21.4 | 77 | 16 | 09KMG7 |
| 12 | 6 | 21.4 | 80 | 2 | 093H79 |
| 13 | 6 | 21.4 | 83 | 2 | 092IK4 |
| 14 | 6 | 21.4 | 123 | 6 | 09GMA9 |
| 15 | 6 | 21.4 | 133 | 11 | 09CS07 |
| 16 | 6 | 21.4 | 136 | 16 | 09A2S3 |

| | | | | | | |
|----|---|------|-----|----|--------|---------------------|
| 17 | 6 | 21.4 | 149 | 6 | 09BGT5 | 09BGT5 macaca fasc |
| 18 | 6 | 21.4 | 150 | 2 | 047303 | 047303 escherichia |
| 19 | 6 | 21.4 | 150 | 2 | 093PM5 | 093PM5 streptococc |
| 20 | 6 | 21.4 | 164 | 11 | 09CQ46 | 09CQ46 mus musculus |
| 21 | 6 | 21.4 | 167 | 2 | 09RKA3 | 09RKA3 streptomyc |
| 22 | 6 | 21.4 | 206 | 15 | 09Q080 | 09Q080 chimpanzee |
| 23 | 6 | 21.4 | 219 | 16 | 09RUJ7 | 09RUJ7 deinococcus |
| 24 | 6 | 21.4 | 231 | 16 | 09CCH2 | 09CCH2 mycobacteri |
| 25 | 6 | 21.4 | 249 | 2 | 09S338 | 09S338 prochloroco |
| 26 | 6 | 21.4 | 262 | 16 | 09SG03 | 09SG03 rhizobium l |
| 27 | 6 | 21.4 | 266 | 2 | 0926G6 | 0926G6 vibrio para |
| 28 | 6 | 21.4 | 269 | 11 | 09DOB9 | 09DOB9 mus musculu |
| 29 | 6 | 21.4 | 275 | 16 | 09A5W0 | 09A5W0 caulobacter |
| 30 | 6 | 21.4 | 293 | 16 | 09V9Y5 | 09V9Y5 staphylococ |
| 31 | 6 | 21.4 | 297 | 2 | 050009 | 050009 mycobacteri |
| 32 | 6 | 21.4 | 300 | 16 | 098C14 | 098C14 rhizobium l |
| 33 | 6 | 21.4 | 303 | 6 | 095KC6 | 095KC6 macaca fasc |
| 34 | 6 | 21.4 | 304 | 16 | 091762 | 091762 pseudomonas |
| 35 | 6 | 21.4 | 306 | 5 | 023151 | 023151 caenorhabdi |
| 36 | 6 | 21.4 | 309 | 16 | 092DK3 | 092DK3 listeria in |
| 37 | 6 | 21.4 | 309 | 16 | 0929K2 | 0929K2 listeria in |
| 38 | 6 | 21.4 | 315 | 4 | 09H2D1 | 09H2D1 homo sapien |
| 39 | 6 | 21.4 | 315 | 4 | 096S07 | 096S07 homo sapien |
| 40 | 6 | 21.4 | 315 | 4 | 096J26 | 096J26 homo sapien |
| 41 | 6 | 21.4 | 315 | 6 | 09BE05 | 09BE05 macaca fasc |
| 42 | 6 | 21.4 | 315 | 6 | 095J75 | 095J75 macaca fasc |
| 43 | 6 | 21.4 | 316 | 17 | 057966 | 057966 pyrococcus |
| 44 | 6 | 21.4 | 318 | 17 | 09V267 | 09V267 pyrococcus |
| 45 | 6 | 21.4 | 327 | 11 | 09DC01 | 09DC01 mus musculu |

ALIGNMENTS

RESULT 1

ID 095KC5 PRELIMINARY; PRT; 501 AA.

AC 095KC5:

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE HYPOTHETICAL 53.4 KDA PROTEIN.

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

OC Cercopithecoidea; Macaca.

OX NCBI_TaxID=9541;

OR

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=MEDULLA OBLONGATA;

RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,

RA Suzuki Y., Sugano S., Hashimoto K.,

RT "Isolation of full-length cDNA clones from macaque brain cDNA

RT libraries."

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB062977; BAB60745.1; -

KW Hypothetical protein.

SQ SEQUENCE 501 AA; 53447 MW; 8C554BBD04EE0470 CRC64;

Query Match 53.6%; Score 15; DB 6; Length 501;
Best local similarity 100.0%; Pred. No. 1.9e-07;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYOGVPRAE 15
Db 242 DFVGEGLYOGVPRAE 256

RESULT 2

ID 096JT2 PRELIMINARY; PRT; 553 AA.

AC 096JT2;

DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE PROSTEIN.
 OC Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=PROSTATE;
 RX MEDLINE=21139094; PubMed=11245466;
 RA Xu J., Kalos M., Stolk J.A., Zaslotti E.J., Zhang X., Houghton R.L.,
 RA Filho A.M., Nolasco M., Badaro R., Reed S.G.;
 RT "Identification and characterization of prostate, a novel prostate-
 RT specific protein."
 RL Cancer Res. 61:1563-1568(2001).
 DR EMBL: AY033593; AKS4386.1;-.
 SQ SEQUENCE 553 AA; 59322 MW; 0AF23FBC742A667 CRC64;

Query Match 53.6%; Score 15; DB 4; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
 ID 294 DFVGEGLYGVPRAE 308
 DB

RESULT 3
 Q95KIT5 PRELIMINARY; PRT; 553 AA.
 AC Q95KIT5
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE HYPOPHYSALIN 59.4 KDA PROTEIN.
 OC Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9541;
 RN
 RP SEQUENCE FROM N.A.
 RC TISSUE=TEMPORAL LOBE RIGHT;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries."
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB060851; BAB46871.1;-.
 KW Hypothetical protein.
 SQ SEQUENCE 553 AA; 59392 MW; 0718F3A91FB3BF1E CRC64;

Query Match 53.6%; Score 15; DB 6; Length 553;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 DFVGEGLYGVPRAE 15
 ID 294 DFVGEGLYGVPRAE 308
 DB

RESULT 4
 Q9VSG8 PRELIMINARY; PRT; 179 AA.
 ID Q9VSG8
 AC Q9VSG8;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
 DE CG1197 PROTEIN.

GN CG1197.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OC NCBI_TaxID=7227;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brannon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Planck C., Baldwin D.,
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., Mosnig A.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshireff A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Sanders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AF003556; AAF50451.1;-.
 DR HSSP: P32889; 1HNR.
 DR FlyBase: FBgn0035866; CG7197.
 DR InterPro: IPR000251; ARF.
 DR InterPro: IPR001806; Ras_trnsfmg.
 DR InterPro: IPR02046; Sar1_GTPBP.
 DR Pfam: PF00025; arf. 1.
 DR PRINTS: PRO0449; RASTRSPRMNG.
 DR PRINTS: PRO0328; SAR1GTPBP.
 DR SMART: SM00177; ARF. 1.
 DR PROSITE: PS01019; ARF. 1.
 KW GTP-binding.
 SQ SEQUENCE 179 AA; 20626 MW; 74DDA1430B5DCDAC CRC64;

Query Match 25.0%; Score 7; DB 5; Length 179;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GEGLYQG 10
 ID 162 GEGLYQG 168
 DB

RESULT 5

```

09R380
ID 09R380 PRELIMINARY; PRT; 364 AA.
AC 09R380
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE GLYCINE BETAININE TRANSPORT ATP-BINDING PROTEIN.
GN OPUAA.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-A3(2);
RA Medienbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,
RA Kinash H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
CC -1 SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
DR EMBL: AL132648; CAB59474.1;
DR InterPro: IPR003593; AAA.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001687; ATP_GTP_A.
DR InterPro: IPR000644; CBS.
DR Pfam: PF00005; ABC_tran. 1.
DR SMART: SM00382; AAA; 1.
DR SMART: SM00116; CBS; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR ATP-binding; Transport.
KW SEQUENCE 364 AA; 39970 MW; B67DE74839610491 CRC64;

Query Match 25.0%; Score 7; DB 2; Length 364;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 9 OGVPRAE 15
DB 134 OGVPRAE 140

```

```

028502
ID 028502 PRELIMINARY; PRT; 569 AA.
AC 028502
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE LONG-CHAIN-FATTY-ACID-COA LIGASE (FADD-7).
GN AF1772.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobales; Archaeoglobaceae;
OC Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RA MEDLINE=98049343; PubMed=9389475;
RA Klenk H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyriades N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirschner E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Cotton M.D., Spriggs T., Artlich P., Kalne B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";
RL Nature 390:364-370(1997).
DR EMBL: AE001112; AAB91290.1;
DR HSSP: P08659; ILCT.
DR TIGR: AF2368;
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP_BINDING; 1.
DR Hypothetical protein; Ligase; Complete proteome.
KW SEQUENCE 562 AA; 62487 MW; CCB8944EFF9DA2B CRC64;

Query Match 25.0%; Score 7; DB 17; Length 562;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 22 HYDEGVR 28
DB 27 HYDEGVR 33

```

Best Local Similarity 100.0%; Pred. No. 30;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 HYDEGVR 28
| | | | |
DB 24 HYDEGVR 30

RESULT 8
Q9RRV6 PRELIMINARY; PRT; 588 AA.

AC Q9RRV6; 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE ABC TRANSPORTER, ATP-BINDING PROTEIN, MSBA FAMILY.
GN DE2379.
CN Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI:
RA MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "genome sequence of the radioresistant bacterium Deinococcus
radiodurans R1.";
RL Science 286:1571-1577(1999).
CC -1 SIMILARITY: BELONGS TO THE ATP-BINDING TRANSPORT PROTEIN FAMILY
(ABC TRANSPORTERS)
DR EMBL: AE002068; AAF1923.1; -
DR TIGR: DR3379; -
DR InterPro: IPR003593; AAA.
DR InterPro: IPR001140; ABC transporter_tmem.
DR InterPro: IPR003439; ABC transporter.
DR InterPro: IPR001360; Glyco_hydro_1.
DR Pfam: PF00064; ABC_membrane_1.
DR Pfam: PF00005; ABC_tran_1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
DR PROSITE: PS00572; GLYCOSYL-HYDROL_FL1; UNKNOWN_1.
DR ATP-binding: Complete proteome; Transport.
KW SEQUENCE 588 AA; 64714 MW; DCB107824C9BF5 CRC64;
SQ

Query Match 25.0%; Score 7; DB 16; Length 588;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 EGTERR 21
| | | | |
DB 229 EGTERR 235

RESULT 9
Q63625 PRELIMINARY; PRT; 1473 AA.

AC Q63625; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE RA9.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

RP SEQUENCE FROM N.A.
RC TISSUE-HIPOCAMPUS; PubMed=8692929;
RX MEDLINE=96293459; PubMed=8692929;
RA Yuvraj A., Paturajan M., Litingtung Y., Joshi R.V., Gentile C.,
RA Gebara M., Corden J.L.;
RT "The C-terminal domain of the largest subunit of RNA polymerase II
interacts with a novel set of serine/arginine-rich proteins.";
RT Proc. Natl. Acad. Sci. U.S.A. 93:6975-6980(1996).
DR EMBL: U49057; AAC52658.1; -
SQ SEQUENCE 1473 AA; 161204 MW; 949E6F5873989BF CRC64;

Query Match 25.0%; Score 7; DB 11; Length 1473;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 GVPRAEG 16
| | | | |
DB 1276 GVPRAEG 1282

RESULT 10
Q69340 PRELIMINARY; PRT; 1958 AA.

AC Q69340; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ORF1, ORF2, AND ORF3.
OS Pseudorabies virus.
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=10345;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=INDIANA-FUNKHAUSER;
RX MEDLINE=91374576; PubMed=1654441.
RA Cheung A.K.;
RT "Cloning of the latency gene and the early protein 0 gene of
pseudorabies virus.";
RT J. Virol. 65:5260-5271(1991).
DR EMBL: M57505; AAA47468.1; -
SQ SEQUENCE 1958 AA; 193474 MW; 0DEC36EDC29E2E68 CRC64;

Query Match 25.0%; Score 7; DB 12; Length 1958;
Best Local Similarity 100.0%; Pred. No. 84;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 RAEGTEA 19
| | | | |
DB 537 RAEGTEA 543

RESULT 11
Q9KMG7 PRELIMINARY; PRT; 77 AA.

AC Q9KMG7; 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN VCA0389.
GN VCA0389.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,

RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,
 RA McDonald L., Uterback T., Fleischmann R.D., Niernan W.C., White O.,
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
 RA Fraser C.M.;
 RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*
 RT *cholerae*.";
 RL Nature 406:477-483(2000).
 DR EMBL: AE004374; AAF96295.1; -
 DR TIGR: VCA0389; -
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 77 AA; 9013 MW; CF2172BF4B38D124 CRC64;

Query Match 21.4%; Score 6; DB 16; Length 77;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 FVGEGL 7
 |||||
 DB 65 FVGEGL 70

RESULT 12
 ID 093H79 PRELIMINARY; PRT; 80 AA.
 AC 093H79;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE FERREDOXINE.
 GN PTEE.
 OS Streptomyces avermitilis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=33903;
 OX [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21477403; PubMed=11572948;
 RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
 RA Shinoze M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
 RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
 RT "Genome sequence of an industrial microorganism *Streptomyces*
 RT *avermitilis*: Deducing the ability of producing secondary
 RT metabolites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
 DR EMBL: AB070949; BAB9311.1; -
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 SQ SEQUENCE 80 AA; 8475 MW; B12A8B1B9D25A556 CRC64;

Query Match 21.4%; Score 6; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 60;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 RAEGTE 18
 |||||
 DB 4 RAEGTE 9

RESULT 13
 ID 09ZIK4 PRELIMINARY; PRT; 83 AA.
 AC 09ZIK4;
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 9.5 KDA PROTEIN (FRAGMENT).
 OS Helicobacter pylori (Campylobacter pylori).
 OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
 OC Helicobacter.
 OX NCBI_TaxID=210;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN-J166;

RX MEDLINE=99007275; PubMed=9789049;
 RA Akopyants N.S., Fradkov A., Diatchenko L., Hill J.E., Siebert P.D.,
 RA Lukyanov S.A., Sverdlov E.D., Berg D.E.;
 RT PCR-based subtractive hybridization and differences in gene content
 RT among strains of *Helicobacter pylori*.
 RL Proc. Natl. Acad. Sci. U.S.A. 95:13108-13113(1998).
 DR EMBL: AF025970; AAC69259.1; -
 DR InterPro: IPR000613; Pseudou_synth.
 DR ProDom: PD001819; Pseudou_synth; 1.
 KW Hypothetical protein.
 FT NON_TER
 SQ SEQUENCE 83 AA; 9503 MW; 795E87B3ED183EFF CRC64;

Query Match 21.4%; Score 6; DB 2; Length 83;
 Best Local Similarity 100.0%; Pred. No. 62;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 VGEGLY 8
 |||||
 DB 24 VGEGLY 29

RESULT 14
 ID 09GMX9 PRELIMINARY; PRT; 123 AA.
 AC 09GMX9;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE HYPOTHETICAL 13.6 KDA PROTEIN.
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OX NCBI_TaxID=9541;
 OX [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN PARIETAL LOBE;
 RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
 RA Suzuki Y., Sugano S., Hashimoto K.;
 RT "Isolation of full-length cDNA clones from macaque brain cDNA
 RT libraries.";
 RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AB047596; BAB12120.1; -
 DR InterPro: IPR001993; Mitoch_carrier.
 DR Pfam: PF00153; mitoch_carrier; 1.
 DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
 KW Hypothetical protein.
 SQ SEQUENCE 123 AA; 13618 MW; 2B553EA863545098 CRC64;

Query Match 21.4%; Score 6; DB 6; Length 123;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 GLYGV 11
 |||||
 DB 80 GLYGV 85

RESULT 15
 ID 09CS07 PRELIMINARY; PRT; 133 AA.
 AC 09CS07;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE 1700073K01RIK PROTEIN (FRAGMENT).
 GN 1700073K01RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYONIC LIVER;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schirini L.M., Staudli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
 Hayashizaki Y.;
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK010880; BAB27243.1; -;
 DR HSSP: P02593; 1CDM.
 DR MGD: MGI:1915476; 1700073K01R1K.
 DR InterPro: IPR002048; EF-hand.
 DR SMART: SM00054; EFP; 2.
 FT NON_TER 133 133
 SQ SEQUENCE 133 AA; 15322 MW; 7D23323D63F7F170 CRC64;

Query Match 21.4%; Score 6; DB 11; Length 133;
 Best Local Similarity 100.0%; Pred. No. 92;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 14 AEGTEA 19
 |||||
 DB 7 AEGTEA 12

Search completed: June 26, 2002, 13:50:42
 Job time: 190 sec